27-я Международная Конференция

DSPA-2025

Доклады

27th International Conference
Digital Signal Processing
and its Applications



СЕРИЯ:
ЦИФРОВАЯ ОБРАБОТКА СИГНАЛОВ
И ЕЁ ПРИМЕНЕНИЕ (выпуск XXVII)

Российское научно-техническое общество радиотехники, электроники и связи им. А.С. Попова

Москва - 2025 г.



XXVII Международная конференция

Цифровая обработка сигналов

и ее применение

DSPA - 2025

26 марта - 28 марта

XXVII International Conference

Digital Signal Processing

and Its Applications

DSPA - 2025

March 26 - March 28

Москва - 2025

27-я Международная конференция «Цифровая обработка сигналов и её применение – DSPA-2025», Москва, Россия, доклады.

Серия: Цифровая обработка сигналов и её применение

Выпуск: XXVII

Рецензенты:

доц. Бабаян П.В., проф. Бартенев В.Г., проф. Брюханов Ю.А., д.т.н. Вашкевич М.И., проф. Витязев В.В., доц. Витязев С.В., проф. Гаврилов К.Ю., prof. Vikram Gadre, доц. Гвоздарёв А.С., член-корр. РАН. Дворкович А.В., д.т.н. Джиган В.И., д.т.н. Зайцев Г.В., к.т.н. Клеопин А.В., доц. Лобов Е.М., д.т.н. Малай И.М., к.т.н. Овинников А.А., доц. Петровский Н.А., д.т.н. Приоров А.Л., проф. Прохоров Ю.Н., проф. Санников В.Г., проф. Сергиенко А.Б., PhD Koushlendra Singh, проф. Сперанский В.С., prof. Radovan Stojanovic, доц. Стротов В.В., prof. Rui Fan, проф. Федоров В.Ф., к.т.н. Ханян Г.С., доц. Хрящев В.В., проф. Цимбал В.А., проф. Чиров Д.С.

© Авторы докладов
© РНТОРЭС имени А.С. Попова

ОРГАНИЗАТОРЫ:

- Российское научно-техническое общество радиотехники, электроники и связи имени А.С. Попова
- Институт радиотехники и электроники имени В.А. Котельникова РАН
- Институт проблем управления имени В.А. Трапезникова РАН
- Компания AUTEX Ltd.

ПРИ УЧАСТИИ:

- Министерство высшего образования и науки РФ
- Федеральное агентство по промышленности РФ
- Институт проблем информатики РАН
- ОАО «РТИ»
- ОАО «ГСКБ «АЛМАЗ-АНТЕЙ»
- ЗАО «Инструментальные системы»
- НТЦ «МОДУЛЬ»
- ОАО «Концерн радиостроения «Вега»
- ОАО «Концерн «СОЗВЕЗДИЕ»
- ГУП НПЦ «Элвис»
- Владимирский государственный университет
- Всероссийский научно-исследовательский институт физико-технических и радиотехнических измерений «ВНИИФТРИ»
- Московский авиационный институт
- Московский государственный технический университет им. Н.Э. Баумана
- Московский технический университет связи и информатики
- Московский Физико-технический институт (университет)
- Московский технологический университет
- Московский институт электронной техники
- Московский энергетический институт (университет)
- Рязанский государственный радиотехнический университет
- Санкт-Петербургский государственный университет телекоммуникаций имени проф. М. А. Бонч-Бруевича
- Санкт-Петербургский государственный электротехнический университет «ЛЭТИ»
- Ульяновский государственный технический университет
- Ярославский государственный университет

МЕЖДУНАРОДНЫЙ ПРОГРАММНЫЙ КОМИТЕТ КОНФЕРЕНЦИИ:

Председатель Международного программного комитета:
академик РАН **Гуляев Юрий Васильевич**
Сопредседатели Международного организационного комитета: академик РАН **Васильев С.Н.**, академик РАН **Каляев И.А.**, **Кирпичников А.П.**, академик РАН **Сигов А.С.**
Председатель Программной группы Международного организационного комитета:
член-корр. РАН **Дворкович Александр Викторович**
Заместитель Председателя Программной группы Международного организационного комитета:
профессор **Витязев Владимир Викторович**

Программная группа Международного организационного комитета:
проф. Аджемов А.С. (Россия), член-корр. РАН Борисов В.И. (Россия), д.т.н. Борисов Ю.И. (Россия), проф. Викрам Гадре (Индия), проф. Михалис Зервакис (Греция), проф. Кушлендра Сингх (Индия), академик РАН Никитов С. А. (Россия), академик РАН Новиков Д.А. (Россия), проф. Радован Стоянович (Черногория), проф. Руи Фан (КНР), академик РАН Чаплыгин Ю.А. (Россия)
Административная группа Международного организационного комитета:
доц. Бабаян П.В., проф. Бартенев В.Г., проф. Брюханов Ю.А., проф. Васильев К.К., доц. Витязев С.В., проф. Гаврилов К.Ю., проф. Джиган В.И., Зубарев Р.В., к.т.н. Клеопин А.В., доц. Лобов Е.М., д.т.н. Малай И.М., д.т.н. Приоров А. Л., проф. Прохоров Ю.Н., доц. Самсонов Г.А., проф. Санников В.Г., доц. Сергиенко А.Б., проф. Сперанский В.С., доц. Стешенко В.Б., проф. Федоров В.Ф., доц. Хрящев В.В., проф. Чиров Д.С.

С О Д Е Р Ж А Н И Е

ПЛЕНАРНЫЕ ДОКЛАДЫ 1

Секция 1. ТЕОРИЯ СИГНАЛОВ И СИСТЕМ 2

Секция 2. ТЕОРИЯ И МЕТОДЫ ЦИФРОВОЙ ОБРАБОТКИ СИГНАЛОВ 41

Секция 3. ОБРАБОТКА СИГНАЛОВ В СИСТЕМАХ ТЕЛЕКОММУНИКАЦИЙ .. 66

Секция 4. ОБРАБОТКА СИГНАЛОВ В РАДИОТЕХНИЧЕСКИХ СИСТЕМАХ... 89

Секция 5. ОБРАБОТКА И ПЕРЕДАЧА ИЗОБРАЖЕНИЙ 101

Секция 6. ОБРАБОТКА СИГНАЛОВ В БИОТЕХНИЧЕСКИХ СИСТЕМАХ..... 119

Секция 7. ПРОЕКТИРОВАНИЕ И ТЕХНИЧЕСКАЯ РЕАЛИЗАЦИЯ СИСТЕМ ЦОС ***..... доклады не представлены***

Секция 8. СИСТЕМЫ «ТЕХНИЧЕСКОГО ЗРЕНИЯ», ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ, РОБОТОТЕХНИКА..... 135

Секция 9. ОБРАБОТКА СИГНАЛОВ В ИЗМЕРИТЕЛЬНЫХ СИСТЕМАХ..... 202



Дворкович Александр Викторович
д.т.н., чл.-корр. РАН, зав. кафедрой МФТИ

Эволюция стандартов цифрового телевидения

Цифровое телевизионное вещание прошло большой путь от первоначальных идей до практического внедрения и дальнейшего развития. Реальное воплощение началось со стандартов первого поколения ATSC, DVB, ISDB в конце 1990-х годов прошлого века. В дальнейшем структура этих систем потребовала развития в связи с ростом новых потребностей телевидения высокой и ультравысокой четкости, а также таких технологий, как повышенная кадровая скорость, расширенная цветовая гамма, высокий динамический диапазон, что привело к возникновению нового поколения стандартов. Это поколение продолжает развертываться в разных странах мира. Новые вызовы требуют дальнейшего развития и совершенствования систем цифрового телевизионного вещания.

Александр Викторович Дворкович, доктор технических наук, член-корреспондент РАН, заведующий кафедрой мультимедийных технологий и телекоммуникаций МФТИ, профессор кафедры радиотехнических приборов и антенных систем МЭИ, лауреат премии Правительства РФ в области науки и техники. Основной сферой его научных интересов является разработка алгоритмов и систем компрессии видео и звука, разработка систем цифровой передачи информации по каналам наземной и спутниковой связи.

Alexander V. Dvorkovich
D. Tech. Sc., Corr. Member of RAS, Head of the Department, MIPT

Evolution of Digital Television Standards

Digital television broadcasting had a long way from the initial ideas to practical implementation and further development. The actual implementation began with the first-generation standards ATSC, DVB, ISDB in the late 1990s of the last century. The structure of these systems subsequently required development in connection with the growth of new needs for high and ultra-high definition television, as well as technologies such as high frame rate, expanded color gamut, and high dynamic range, which led to the emergence of a new generation of standards. This generation continues to be deployed in different countries around the world. New challenges require further development and improvement of digital television broadcasting systems.

Alexander V. Dvorkovich, Doctor of Technical Sciences, Corresponding Member of the Russian Academy of Sciences, Head of the Department of Multimedia Technologies and Telecommunications at MIPT, Professor of the Department of Radio Engineering Devices and Antenna Systems at MPEI, Laureate of the Russian Federation Government Prize in Science and Technology. The main area of his scientific interests is the development of algorithms and systems for video and audio compression, the development of systems for digital transmission of information via terrestrial and satellite communication channels.

ИССЛЕДОВАНИЕ ИСКАЖЕНИЙ ШИРОКОПОЛОСНЫХ СИГНАЛОВ ИЗ-ЗА ЧАСТОТНОЙ ДИСПЕРСИИ В ТРАНСИОНОСФЕРНЫХ РАДИОКАНАЛАХ

студ. Цой Д. Ч.[1]

[1] Ордена Трудового Красного Знамени федеральное государственное бюджетное образовательное учреждение высшего образования «Московский технический университет связи и информатики»

Исследовано влияние дисперсии трансионосферного радиоканала на широкополосный сигнал. Проведено сравнение откликов согласованного фильтра на искаженный и неискаженный широкополосный сигнал, прошедший трансионосферный радиоканал с целью оценки влияния изменения параметров частотной дисперсии и окружающей среды на принимаемый сигнал. Получены графики зависимостей отношения максимумов амплитуд сигнала искаженного к сигналу неискаженному от изменения параметров частотной дисперсии ионосферы, различных времен года и разного времени суток.

Введение

В современных радиотехнических системах, использующих трансионосферный радиоканал, для обеспечения требуемой скорости передачи данных применяется расширение полосы частот сигналов. Увеличение полосы пропускания приводит к возникновению дисперсионных искажений, обусловленных различием фазовых скоростей спектральных компонент широкополосного сигнала. Изучение механизмов дисперсии и разработка моделей для их описания являются актуальными задачами, значимыми как для фундаментальных исследований в радиофизике, так и для прикладных решений в областях связи и навигации. Предлагаемая работа посвящена исследованию изменения характеристик частотной дисперсии в трансионосферных радиоканалах в зависимости от сезона года и времени суток и ее влияния на искажение широкополосных сигналов.

Цель работы – анализ влияния частотной дисперсии в трансионосферных радиоканалах с оценкой значимости параметров дисперсии различных порядков в искажениях широкополосных сигналов.

Метод, используемый для оценки влияния дисперсии трансионосферного радиоканала на широкополосный сигнал

В основе исследования лежит сопоставление выходных сигналов согласованного фильтра, полученных для искаженного и исходного сигналов. Моделирование фазовых искажений, привнесенных ионосферой, производилось в частотной области с использованием параметров ионосферы, полученных в результате вычислительного эксперимента.

В данной модели задача распространения описывается частотной характеристикой (ЧХ) [1] эквивалентной среде радиоканала. При этом частотная характеристика канала имеет динамический характер и определяется выражением:

$$H(j\omega,t) = H_0(\omega,t)\exp\left[-j\varphi(\omega,t)\right], \quad (1)$$

где $H_0(\omega,t)$ – амплитудно-частотная характеристика (АЧХ) канала; $\varphi(\omega,t)$ – фазо-частотная (ФЧХ) канала.

ФЧХ канала можно представить в виде степенного ряда Тейлора до третьего порядка [2]:

$$\begin{aligned}\varphi(\omega,t) &= \varphi(\bar{\omega},t) + \varphi'(\bar{\omega},t)\cdot\Omega + \varphi''(\bar{\omega},t)\cdot\frac{\Omega^2}{2!} + \varphi'''(\bar{\omega},t)\cdot\frac{\Omega^3}{3!} + \ldots \approx \\ &\approx \varphi(\bar{f},t) + 2\pi\tau_g(\bar{f},t)\cdot F + \pi s(\bar{f},t)\cdot F^2 + \frac{\pi}{2}v(\bar{f},t)\cdot F^3,\end{aligned} \quad (2)$$

где $F = \frac{\Omega}{2\pi} = f - \bar{f}$ – разностная частота; $\tau_g(\bar{f},t) = d\varphi / d\omega$, $s(\bar{f},t) = d\tau_g / df = 2\pi\cdot\varphi''$, $v(\bar{f},t) = d\tau_g{}^2 / df^2 = 2\pi\cdot\varphi'''$ – параметры фазовой дисперсии первого, второго и третьего порядков.

Полную фазу можно разделить на линейную $\varphi_{л}\left(\bar{f},t\right)=\varphi\left(\bar{f},t\right)+2\pi\tau_{g}(\bar{f},t)\cdot F$ и нелинейную $\varphi_{н}(\bar{f},t)=\pi s(\bar{f},t)\cdot F^{2}+\frac{\pi}{2}v(\bar{f},t)\cdot F^{3}$ составляющие.

Дисперсионная характеристика (ДХ) представляет из себя функцию вида:

$$\tau_{g}(F)=d\varphi/d\omega\approx\tau_{g}+s(F)+\frac{v}{2}(F)^{2}. \quad (3)$$

Чтобы получить общее решение, пригодное для практического использования в расчетах параметров дисперсии, были выполнены вычисления для опорной частоты f_r. В качестве опорной была выбрана частота $f_r=1$ГГц. Выражения для параметров дисперсии приведены к виду:

$$\tau(\hat{f})=\int_{s}\frac{dz}{c}+\frac{k\cdot N_t}{2\cdot c\cdot f_r^{2}}\cdot\frac{1}{\hat{f}^{2}}=\int_{s}\frac{dz}{c}+\frac{\tau_k}{\hat{f}^{2}}; \quad (4)$$

$$s(\hat{f})=\frac{k\cdot N_t}{c\cdot f_r^{3}}\cdot\frac{1}{\hat{f}^{3}}=\frac{s_k}{\hat{f}^{3}}; \quad (5)$$

$$v(\hat{f})=\frac{3k\cdot N_t}{c\cdot f_r^{4}}\cdot\frac{1}{\hat{f}^{4}}=\frac{v_k}{\hat{f}^{4}};\ , \quad (6)$$

где $\hat{f}=\bar{f}/f_r$ – относительная частота; f_r – опорная частота; $\bar{f}$ – заданная частота; τ_k – дисперсионная характеристика (ДХ) на эталонной частоте; s_k – параметр дисперсии второго порядка на опорной частоте; v_k – параметр дисперсии третьего порядка на опорной частоте; $k=80.5[\text{м}^3/\text{с}^2]$; $N_t=\int_0^z N(z)dz$ – при больших значениях z полное электронное содержание (ПЭС) ионосферы [3].

Результаты моделирования

Исследование проводилось в пакете прикладных программ MATLAB. В модели были заданы следующие параметры: ширина спектра 220 МГц, символьной скорости 200 МБод с учетом параметров частотной дисперсии до третьего порядка.

На рисунках 1-2 для примера приведены результаты сравнения откликов согласованного с широкополосным сигналом фильтра на неискаженный и искаженный сигналы для разного времени суток при соответствующих значениях параметров нелинейной частотной дисперсии для заданной частоты. Для остальных времен года исследование проводилось аналогичным методом.

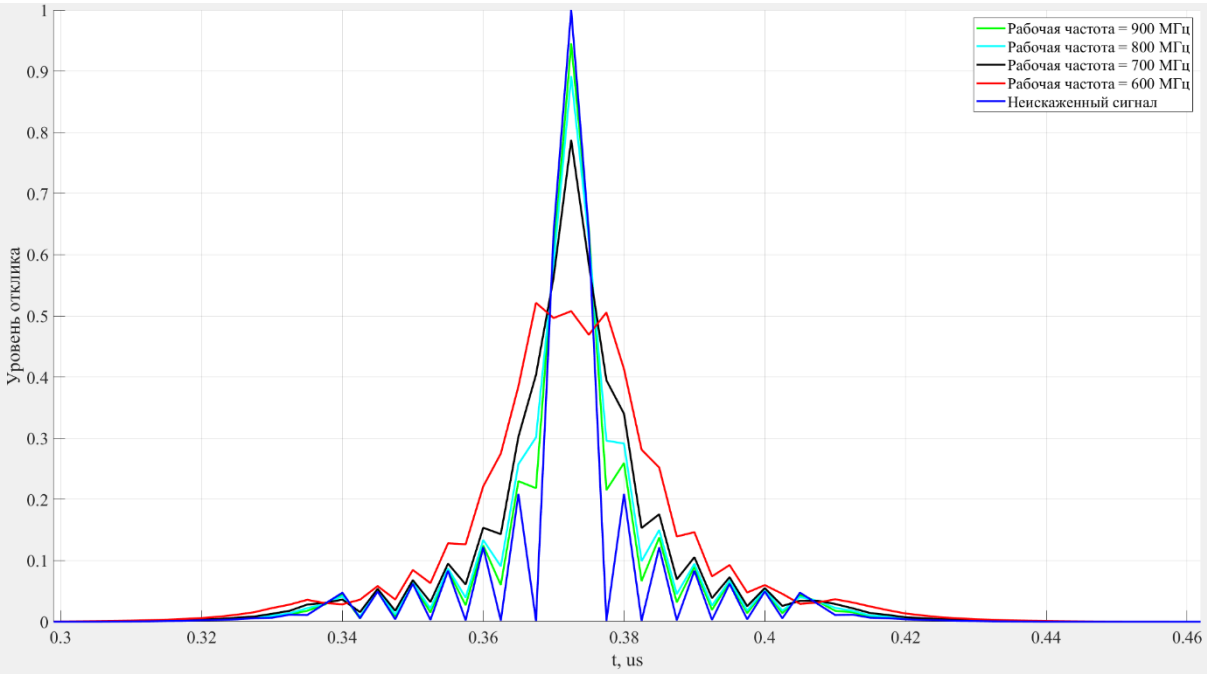


Рисунок 1 – Отклики СФ на искаженный и неискаженный сигнал при прохождении через трансионосферный канал (лето(июнь), день; $\overline{f} = 600\,\text{МГц}$, $s = -1.2 \cdot 10^{-4}$ мкс/МГц, $v = 1.68 \cdot 10^{-7}$ мкс/МГц^2; $\overline{f} = 700\,\text{МГц}$, $s = -0.72 \cdot 10^{-4}$ мкс/МГц, $v = 1.44 \cdot 10^{-7}$ мкс/МГц^2; $\overline{f} = 800\,\text{МГц}$, $s = -0.5 \cdot 10^{-4}$ мкс/МГц, $v = 1.27 \cdot 10^{-7}$ мкс/МГц^2; $\overline{f} = 900\,\text{МГц}$, $s = -0.35 \cdot 10^{-4}$ мкс/МГц, $v = 1.12 \cdot 10^{-7}$ мкс/МГц^2)

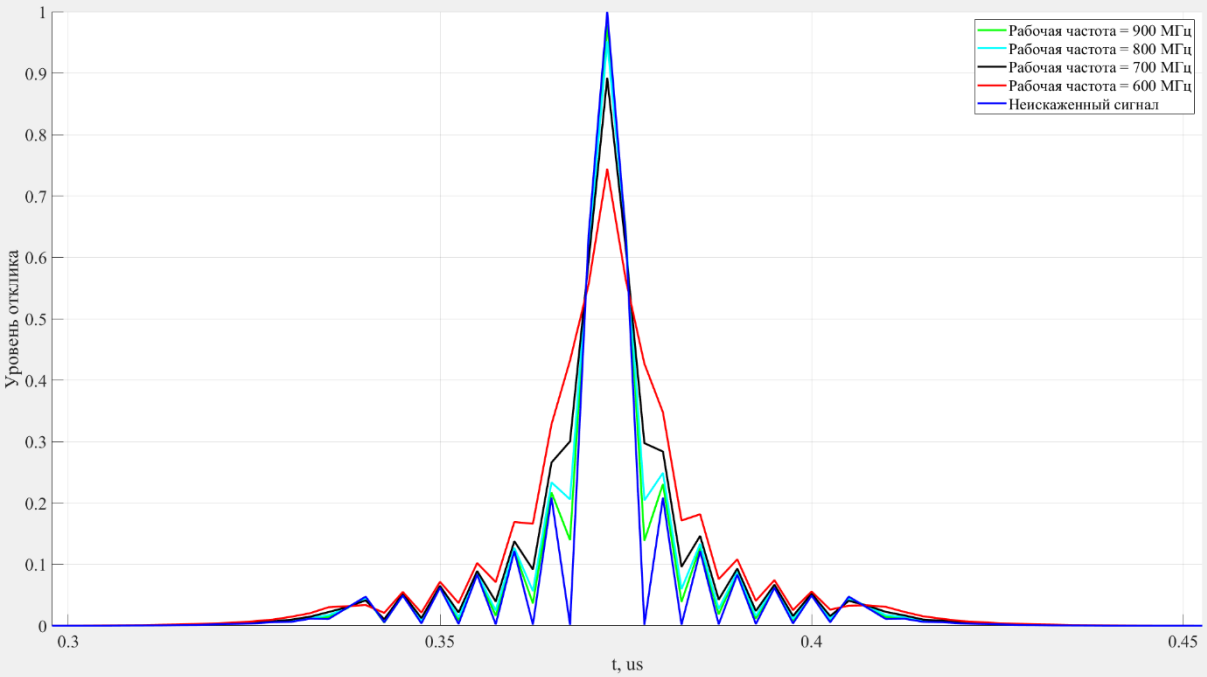


Рисунок 2 – Отклики СФ на искаженный и неискаженный сигнал при прохождении через трансионосферный канал (лето(июнь), ночь; $\overline{f} = 600\,\text{МГц}$, $s = -0.8 \cdot 10^{-4}$ мкс/МГц, $v = 7.74 \cdot 10^{-8}$ мкс/МГц^2; $\overline{f} = 700\,\text{МГц}$, $s = -0.5 \cdot 10^{-4}$ мкс/МГц, $v = 6.63 \cdot 10^{-8}$ мкс/МГц^2; $\overline{f} = 800\,\text{МГц}$, $s = -0.33 \cdot 10^{-4}$ мкс/МГц, $v = 5.804 \cdot 10^{-8}$ мкс/МГц^2; $\overline{f} = 900\,\text{МГц}$, $s = -0.22 \cdot 10^{-4}$ мкс/МГц, $v = 5.16 \cdot 10^{-8}$ мкс/МГц^2)

Из этих данных можно прийти к выводу о зависимости отношений максимальных значений амплитуд сигнала искаженного к сигналу неискаженному от изменения параметров частотной дисперсии ионосферы, времени года и суток. На рисунках 3-4 изображены графики изменения отношений амплитуд сигналов в зависимости от изменения заданной частоты $\overline{f}$.

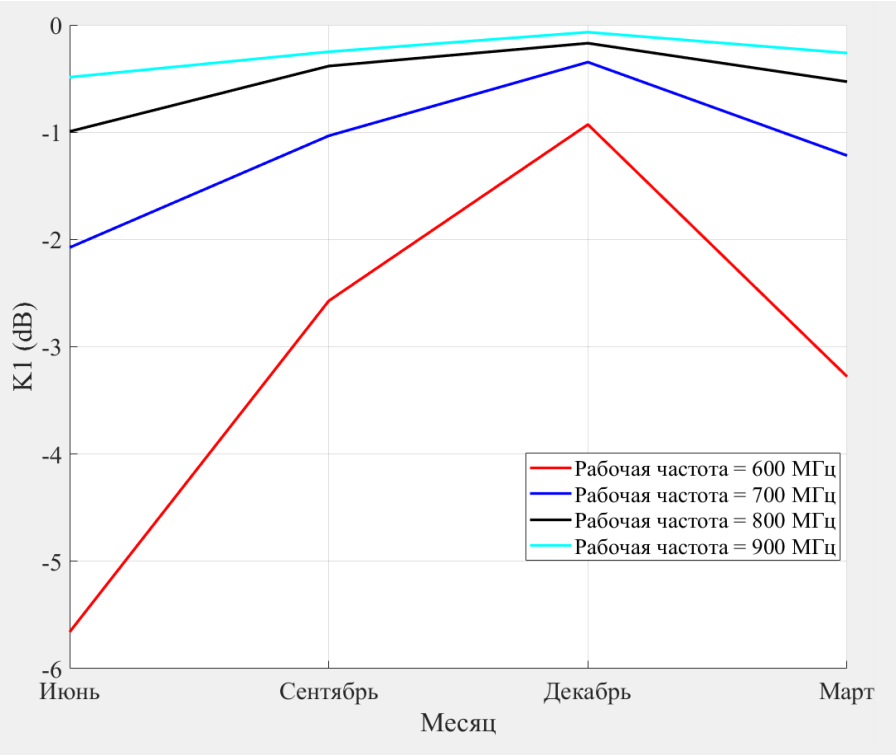


Рисунок 3 – График изменения отношений амплитуд сигналов на различных рабочих частотах в дневное время суток

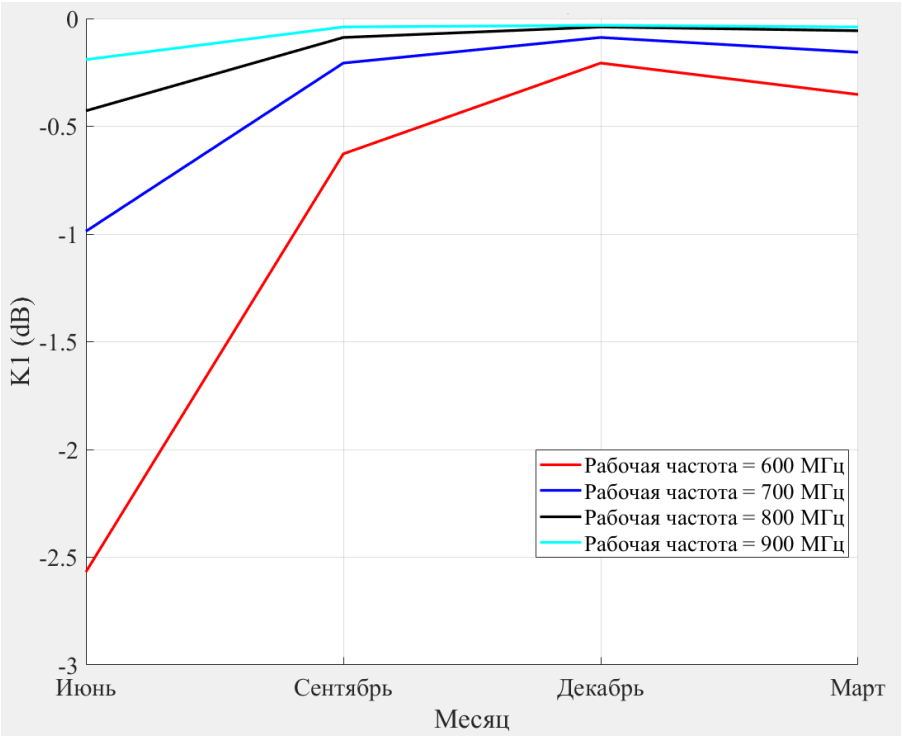


Рисунок 4 – График изменения отношений амплитуд сигналов на различных рабочих частотах в ночное время суток

Заключение

В данной работе проведено исследование влияния частотной дисперсии трансионосферных радиоканалов на широкополосные сигналы. Представленный метод оценки дисперсионных искажений основан на сравнении откликов согласованного фильтра на искаженный и неискаженный сигналы. Были получены данные, демонстрирующие изменение отношений максимальных амплитуд искаженного под воздействием среды и неискаженного сигналов в зависимости от рабочей частоты сигнала, сезона и времени суток. Согласно полученным результатам, наименьшее затухание сигнала наблюдалось в зимний период ночное время суток на заданной частоте $\overline{f} = 900\,\text{МГц}$ и был равен -0.0393 дБ.

Литература

1. Д. В. Иванов, В. А. Иванов, Н. В. Рябова, М. И. Рябова, А. А. Кислицын, А. А. Чернов, Н. А. Конкин. Анализ математических приближений для оценки частотной фазовой дисперсии широкополосных трансионосферных каналов связи. Картографирование полосы когерентности // Материалы Всероссийской научной конференции. 2017. С. 315-322.
2. А. А. Кислицын. Комплекс адаптивной компенсации энергетических потерь сигналов из-за частотной дисперсии в трансионосферных радиоканалах систем спутниковой связи // Системы, сети и устройства телекоммуникаций: дис. … канд. тех. наук : 05.12.13/ Кислицын Алексей Александрович; науч. рук. Н.В.Рябова; МТУСИ — Москва, 2020. — 157 с.
3. В. А. Иванов, Н. В. Рябова, М. И. Рябова, А. А. Кислицын. Влияние частотной дисперсии на распространение сигналов при трансионосферном зондировании // Материалы Международных научно-технических конференций. Т. 2. 2014. С. 136-139.

RESEARCH OF DISTORTIONS OF BROADBAND SIGNALS DUE TO FREQUENCY DISPERSION IN TRANSPHERE RADIO CHANNELS

Tsoy D.Ch.[1]

1 Moscow Technical University of Communication and Informatics (MTUCI)

A method for evaluating the effect of transionospheric radio channel dispersion on a broadband signal was described. A comparison of the matched filter response to the corrupted and non-corrupted broadband signal that passed the transionospheric radio channel to assess the effect of changes in frequency dispersion parameters and the environment on the received signal was carried out. The figures of dependences of maximum amplitudes of the corrupted signal amplitudes to the non-corrupted signal amplitudes from changes in frequency dispersion parameters of the ionosphere, different seasons and different times of day were obtained.

МОДЕЛИРОВАНИЕ ПЕРЕДАЧИ ИНФОРМАЦИОННЫХ СООБЩЕНИЙ В БЕСПРОВОДНЫХ СЕТЯХ

проф. Преображенский А.П.[1], преп. Аветисян Т.В.[2], студ. Дубовик К.А.[1], Маркин В.В.[1]

[1] Воронежский институт высоких технологий
[2] Колледж Воронежского институт высоких технологий

В работе обсуждаются возможности решения задачи, связанной с передачей информационных сообщений в беспроводных сетях. Рассмотрен пример анализируемой ячеистой сетевой структуры. Рассмотрена модель передачи сигналов разных видов по каналам с низкой и высокой скоростью. Проведен анализ соответствующих потерь. Для различных вариантов передачи информации осуществлялись процессы имитационного моделирования. Проведено исследование того, как число узлов в анализируемой ячеистой сетевой структуре оказывает влияние на характеристики передаваемых информационных сообщений.

В настоящее время можно наблюдать реализацию цифровой трансформации в различных прикладных областях. Происходит персонализация телекоммуникационных услуг, глобализация обработки данных и конвергенция телекоммуникационных сетей. конвергенцией сетей, глобализацией и персонализацией телекоммуникационных услуг. При создании современных средств передачи данных разработчики ориентируются на комплексные решения, объединяющие в себе несколько технологий. За счет этого могут поддерживаться услуги по передаче трафика разного типа. Объемы передачи трафика во всем мире возрастают с каждым годом. В связи с тем, что мобильность абонентов увеличивается, наблюдается развитие беспроводных технологий. При этом одной из важных задач является обеспечение бесперебойного доступа к каналу связи [1]. Отдельно следует отметить mesh-сети. В них предусмотрены возможности для того, чтобы каждые устройства работали в виде маршрутизатора, который будет соответствующим образом осуществлять направление трафика.

Использование указанной технологии обеспечивает. С одной стороны – экономичность, а с другой стороны – весьма высокую устойчивость в работе. Будет обеспечиваться перенаправление по другим маршрутам, если какие-то из устройств перестанут работать. То есть, наблюдаются такие характеристики, как самоадаптация и самовосстановление. Сама создаваемая система передачи информационных сообщений достаточно проста в обслуживании, удобным образом развертывается, оперативным образом выполняет возложенные на нее функции. То, насколько mesh-сети будут эффективными и производительными определяется такими параметрами, как время ожидания ответа, число узлов, возможных ошибок в ходе передачи информации [2]. С точки зрения практического использования можно отметить, что поскольку частота, на которой работает оборудование, является высокой, будет хорошая устойчивость относительно существующих помех. Кроме того, наблюдается высокое проникновение сигналов через разные препятствия в системе mesh-сетей. Необходимо учитывать, что в ряде случаев может потребоваться больше энергии для функционирования системы, а также обеспечение повторной передачи данных вследствие того, что число ошибок растет по мере того, как будет большее число узлов [3]. По характеристикам энергопотребления и числу узлов важно стремиться к некоторому балансу. Для того, чтобы быть уверенным в стабильной работе mesh-сетей [4], следует ориентироваться на результаты предварительным образом проведенного моделирования.

Проведем анализ сети с ячеистой структурой на основе инструментального средства Scilab. На рис. 1 приведен пример анализируемой сетевой структуры. Пусть рассматривается ситуация, когда одновременным образом происходит передача трафика разного типа, например, речевые сигналы и видеосигналы.

Показатель использования канала при передаче речевых сигналов:

$$\rho_1 = \frac{\lambda_1}{\mu_1}, b_1 = 1, \quad (1)$$

здесь λ_1 – показывает, с какой интенсивностью будут поступать заявки; μ_1 – показывает, с какой интенсивностью будут обрабатываться заявки; b_1 – характеристика скорости движения заявок; n_1 – общее число заявок.

Показатель использования канала при передаче видеосигналов:

$$\rho_2 = \frac{\lambda_2}{\mu_2}, b_2 = 1, \quad (2)$$

здесь λ_2 – показывает, с какой интенсивностью будут поступать заявки; μ_2 – показывает, с какой интенсивностью будут обрабатываться заявки; b_2 – характеристика скорости движения заявок; n_2 – общее число заявок.

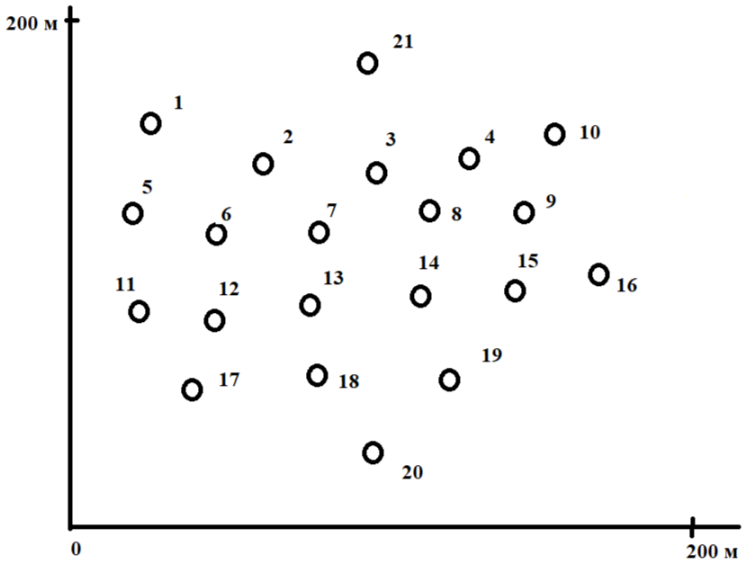


Рисунок 1 – Схема компьютерной сети с ячеистой структурой

Будем считать, что анализируемая информационно-телекоммуникационная система может быть в некоторых состояниях (n_1, n_2). Эти состояния находятся в соответствующем пространстве:

$X = \{(n_1, n_2): n_1 \geq 0, n_2 \geq 0, n_1 b_1 + n_2 b_2 \leq C\}$.

Подпространства для высокоскоростных и низкоскоростных каналов с учетом возможных отказов в передаче данных, определяются следующим образом:

Высокоскоростной канал:

$$X_2 = \{(n_1, n_2): n_1 b_1 + n_2 b_2 \leq C - b_2\}. \quad (3)$$

Условие отказа:

$$\overline{X} = X \setminus X_2 = \{(n_1, n_2): n_1 b_1 + n_2 b_2 > C - b_2\}. \quad (4)$$

Низкоскоростной канал:

$$X_1 = \{(n_1, n_2): n_1 b_1 + n_2 b_2 \leq C - b_1\}, \quad (5)$$

Условие отказа:

$$\overline{X} = X \setminus X_1 = \{(n_1, n_2): n_1 b_1 + n_2 b_2 > C - b_1\}. \quad (6)$$

В том случае, если в одном канале разные составляющие трафика будут перемещаться с разной скоростью можно проводить анализ такого пространства состояний:

$X = \{(n_1, n_2): N_1 \geq n_1 \geq 0, n_2 \geq 0, n_1 b_1 + n_2 b_2 \leq C\}$. Чтобы определить вероятность попадания в состояния, следует использовать такое выражение:

$$p(n_1,n_2)=\frac{\rho_1^{n_2}}{n_2!}\frac{\rho_1^{n_1}}{n_1!}\left[\sum_{n_1=0}^{\left[\frac{c}{N_1}\right]}\sum_{n_2=0}^{\left[\frac{c-n_1N_1}{b_2}\right]}\frac{\rho_1^{n_1}}{n_1!}\frac{\rho_1^{n_2}}{n_2!}\right]^{-1}. \quad (7)$$

В результате рассмотрения было установлено, что низкоскоростные соединения составляли от 10% до 90%. Показатель использования канала для одного узла менялся в пределах 0,75 - 0,95. Значения потерь для разных каналов указаны в табл. 1. Проведенный анализ продемонстрировал, что чем больше скорость будет у заявок, тем большие потери будут наблюдаться.

Таблица 1. Анализ потерь, соответствующих низкой и высокой скорости

Доля высокоскоростных соединений	**Использование высокоскоростного канала**				**Использование низкокоскоростного канала**			
	ρ = 0,76	ρ = 0,81	ρ = 0,91	ρ = 0,96	ρ = 0,76	ρ = 0,81	ρ = 0,86	ρ = 0,96
0,1	0,0000	0,0356	0,155	0,358	0,0003	0,0005	0,0008	0,0092
…	…	…	…	…	…	…	…	…
0,8	0,0661	0,0855	0,153	0,178	0,0008	0,0022	0,0042	0,0052

С учетом разных вариантов формирования сетевых структур осуществлялись процессы имитационного моделирования. Учитывалось, что будут выполняться такие начальные установки:

1. Ретрансляция ведется на базе узлов 3, 6, 7, 8, 9, 12, 13, 14, 15, 17;
2. Длительность цикла для моделирования соответствует 4 секундам.

Были рассмотрены разные возможные сценарии с учетом того, что используется распределитель позиций узлов (NPA):

Вариант 1:

1. От одного узла максимум может поступать 5 пакетов.
2. В узел 9 от узла 11 будет поступление пакетов, имеющих размер 36 байт, при скорости 2 пакета в секунду.

Вариант 2:

1. От одного узла максимум может поступать 10 пакетов.
2. В узел 4 от узла 5 будет поступление пакетов, имеющих размер 36 байт, при скорости 3 пакета в секунду.

Вариант 3:

1. От одного узла максимум может поступать 15 пакетов.
2. В узел 10 от узла 1 будет поступление пакетов, имеющих размер 36 байт, при скорости 2 пакета в секунду.
3. В узел 9 от узла 11 будет поступление пакетов, имеющих размер 36 байт, при скорости 3 пакета в секунду.
4. В узел 4 от узла 5 будет поступление пакетов, имеющих размер 36 байт, при скорости 4 пакета в секунду.

Вариант №4:

1. От одного узла максимум может поступать 15 пакетов.
2. В узел 10 от узла 1 будет поступление пакетов, имеющих размер 36 байт, при скорости 2 пакета в секунду.
3. В узел 16 от узла 21 будет поступление пакетов, имеющих размер 36 байт, при скорости 3 пакета в секунду.
4. В узел 4 от узла 5 будет поступление пакетов, имеющих размер 36 байт, при скорости 4 пакета в секунду.
5. В узел 9 от узла 11 будет поступление пакетов, имеющих размер 36 байт, при скорости 5 пакетов в секунду.

На рис. 2 показана зависимость бит, которые передаются в сетевой структуре от числа узлов в ней. Монотонная зависимость на графике связана с тем, что через промежуточные узлы проходит трафик и требуется, чтобы дополнительные сообщения были отправлены.

На рис. 3 показано, как зависят число коллизий от числа узлов в сетевой структуре. Монотонная зависимость на графике связана с тем, что по мере увеличения числа сообщений коллизий будет больше.

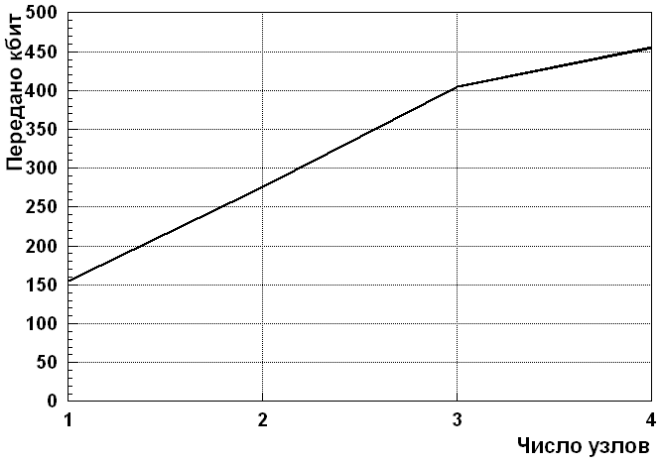


Рисунок 2 –Зависимость бит, которые передаются в сетевой структуре от числа узлов в ней.

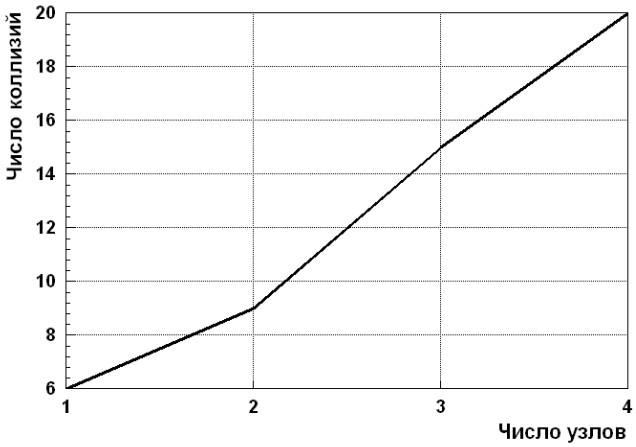


Рисунок 3 –Зависимость числа коллизий от числа узлов в сетевой структуре

На рис. 4 показано, как влияет число узлов, которые передают информацию, на число ошибок (число ошибочно обработанных пакетов в единицу времени). Целостность сообщений будет нарушаться нелинейным образом по мере роста числа узлов.

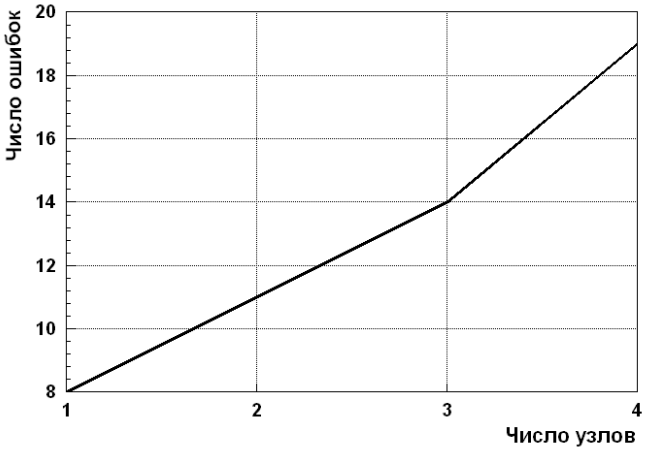


Рисунок 4 – Зависимость числа ошибок от количества передающих узлов

На рис. 5 показано, как влияет число узлов на отправляемые сообщения. Анализ показал, что зависимость является нелинейной.

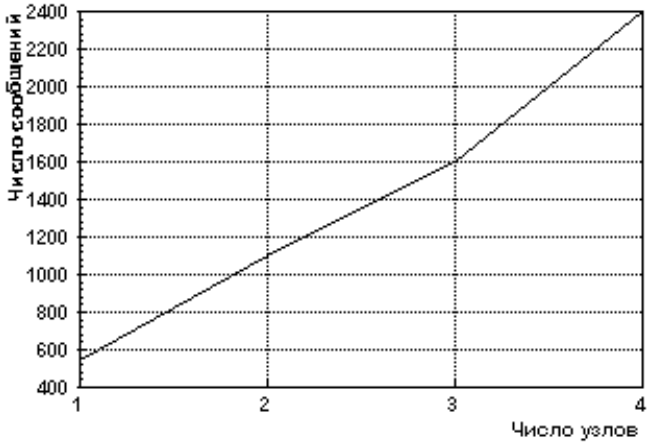


Рисунок 5 – Зависимость сообщений от числа узлов

Вывод. Исследованы возможности передачи в ячеистой компьютерной сети. Процессы моделирования проводились в Scilab. Рассмотрены различные варианты передачи информации. Исследованы зависимости между такими параметрами как число узлов, сообщений, бит, а также возможных ошибок.

Литература

1. А.В.Захаров, С.В.Чаенко, А.А.Демидов, К.А. Кривулин. Сравнительный анализ mesh wi-fi сетей и традиционных wi-fi сетей // Матрица научного познания. 2023. № 7-2. С. 58-63.
2. В.А. Маленко, Е.А.Редькина. Применение беспроводной wi-fi mesh-сети для передачи данных в системе хранения // Современные проблемы радиоэлектроники и телекоммуникаций. 2023. № 6. С. 85.
3. А.В. Жариков, О.В. Константинова, О.С. Терновой. Построение модели mesh-сети с кэшированием трафика на основе p2p-механизма // Динамика систем, механизмов и машин. 2021. Т. 9. № 3. С. 110-116.
4. И.В. Родыгина, В.А. Новак. Моделирование беспроводной mesh-сеть на основе спецификации ZIGBEE // Известия ЮФУ. Технические науки. 2021. № 5 (222). С. 60-68.

MODELING OF INFORMATION MESSAGE TRANSMISSION IN WIRELESS NETWORKS

Preobrazhenskiy A.P.[1], Avetisyan T.V.[2], Dubovik K.A.[2], Markin V.V.[2]

[1] Voronezh institute of high technologies
[2] College of Voronezh institute of high technologies

The paper discusses the possibilities of solving the problem related to the transmission of information messages in wireless networks. An example of the analyzed cellular network structure is considered. A model of transmission of signals of different types through channels with low and high speed is considered. An analysis of the relevant losses was carried out. Simulation modeling processes were carried out for various options for transmitting information. A study is carried out on how the number of nodes in the analyzed mesh network structure affects the characteristics of transmitted information messages.

ОТНОШЕНИЕ СИГНАЛ/ШУМ ПРИ КВАЗИОПТИМАЛЬНОМ ПРИЕМЕ ШИРОКОПОЛОСНЫХ СИГНАЛОВ С ПРИМЕНЕНИЕМ УСТРОЙСТВА ПОДАВЛЕНИЯ ПОМЕХ НА БАЗЕ БАНКА ЦИФРОВЫХ ФИЛЬТРОВ

доц. Лобов Е.М., асс. Шубин Д.Н., студ. Григорьева А.Д.

Московский технический университет связи и информатики

В работе получены выражения для расчета отношения сигнал/шум при квазиоптимальном приеме широкополосных сигналов с применением устройства подавления помех на базе банка цифровых фильтров. С использованием записей эфира проведены расчеты отношения сигнал/шум при приеме сигнала в полосе 160 кГц в условиях узкополосных помех. Расчетным путем показано, что разница между потенциально достижимым значением отношения сигнал/шум и достигаемом значением отношения сигнал/шум при использовании устройства подавления помех на базе банка цифровых фильтров с 256-ю фильтрами составляет менее 0.1 дБ для обработанных записей.

Введение

Обработка широкополосного сигнала в декаметровом диапазоне длин волн осуществляется в условиях наличия множества узкополосных помех, количество которых заранее неизвестно, ровно, как и их структура и характеристики. В связи с чем, компенсационные методы учитывающие структуру помехи не всегда могут быть применены на практике, особенно, при наличии множества разнотипных помех. Методы диаграммообразования и формирования нуля диаграммы направленности антенны в направлении помехи не применимы при наличии секторных крупногабаритных антенн. Алгоритмы адаптации по форме полезного сигнала также не применимы в условиях низкой энергетики принимаемого широкополосного сигнала.

Наиболее простым вариантом борьбы с узкополосными помехами является режекция части спектра, пораженной помехами, в частотной области на базе динамически перестраиваемого фильтра-режектора, преобразования Фурье или банка цифровых фильтров.

Однако, при наличии множества узкополосных помех и шума, совокупное аддитивное мешающее воздействие в силу центральной предельной теоремы можно считать гауссовским случайным процессом с неравномерной спектральной плотностью мощности, т.е. окрашенным гауссовским шумом. Метод оптимального приема в условиях окрашенного гауссовского шума известен и состоит в использовании предварительного фильтра, частотная характеристика которого обратно пропорциональна спектральной плотности мощности помехи [1, 2].

В [3-5] предложено устройство подавления помех на базе банка цифровых фильтров, которое осуществляет разделение широкополосного сигнала на узкополосные компоненты, измерения средней мощности узкополосных компонент, и нормирование на измеренное значение и восстановление широкополосного сигнала. Указанное устройство представляет собой вариант технической реализации фильтра предварительной обработки с возможностью автоматической подстройки к динамически изменяющимся помеховым условиям в эфире. Установлено, что применение устройства позволяет проводить квазиоптимальный прием в условиях окрашенного гауссовского шума и показано (в результате обработки множества сигналов эфира), что достигается выигрыш до 2 дБ в отношении сигнал/шум по сравнению с режекцией помех. Результаты экспериментов также позволили сформулировать рекомендации по выбору числа каналов банка цифровых фильтров.

Несмотря на упомянутый выигрыш при применении устройства подавления помех на базе банка цифровых фильтров, недостаточно ясно, насколько получаемый на практике уровень отношения сигнал/шум близок к потенциально достижимому при оптимальном приеме. Чтобы оценить близость квазиоптимальной обработки к оптимальной требуется получить аналитические выражения для расчета отношения сигнал/шум при использовании устройства подавления банка цифровых фильтров, провести расчеты и сравнить с потенциально достижимым значением.

Таким образом целью работы является расчет отношения сигнал/шум при квазиоптимальном приеме широкополосных сигналов с применением устройства подавления помех на базе банка цифровых фильтров и сравнение полученного значения с потенциально достижимым.

Устройство подавления помех на базе банка цифровых фильтров

На рисунке 1 представлена структурная устройства подавления помех на базе банка цифровых фильтров с блоком субполосной обработки на низкой частоте дискретизации. Банк фильтров состоит из [6-8]: $2M$ фильтров анализа $\dot{H}_k(z)$, $k=0\div 2M-1$, $2M$ дециматоров, блока обработки сигналов на низкой частоте дискретизации, $2M$ интерполяторов, $2M$ фильтров синтеза $\dot{F}_k(z)$, $k=0\div 2M-1$, блок суммирования. Длина импульсной характеристики всех фильтров одинаковая и равна N отсчетов. Фильтры анализа и синтеза рассчитываются на основе одного и того же низкочастотоного фильтра-прототипа, путем смещения его комплексной частотной характеристики в частотной области. Поэтому амплитудно-частотные характеристики (АЧХ) фильтров анализа и синтеза с равными индексами одинаковые. Банки цифровых фильтров нашли широкое применение в обработке широкополосных сигналов декаметрового диапазона частот (см., например, [9-12]).

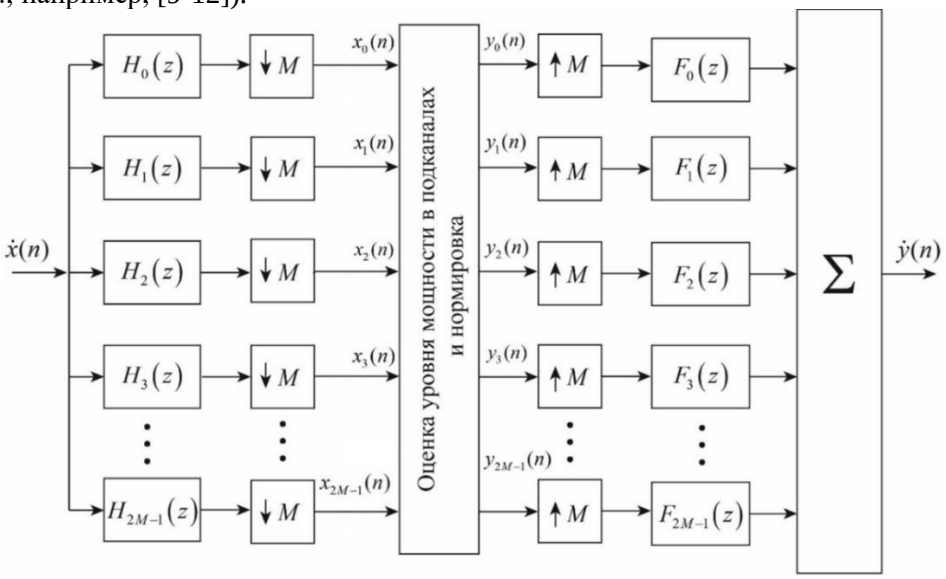


Рисунок 1 – Структура устройства подавления помех

В качестве блока обработки сигналов на низкой частоте дискретизации в каждом узкополосном канале используется блок оценивания средней мощности сигнала на временном окне заданной длительности (L отсчетов) и нормирования сигнала на указанное значение (из [1, 2], см. рисунок 2). Другими словами, узкополосный сигнал в каждом низкочастотном канале умножается на коэффициент c_k, обратно пропорциональный средне мощности сигнала.

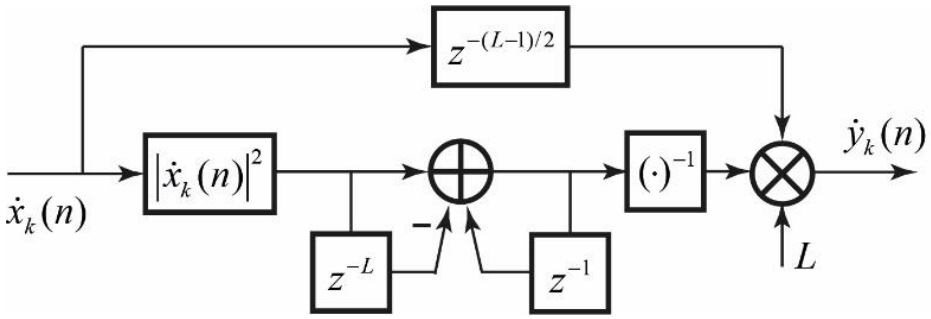


Рисунок 2 – Оценивание средней мощности сигнала и нормирование сигнала на оцененное значение

Аналитические выражения для вычисления отношения сигнал/шум

Спектральная плотность сигнала, прошедшего банк фильтров, имеет вид

$$\dot{Y}_{BF}(j\omega)\approx\sum_{k=0}^{2M-1}\dot{H}_k(j\omega)c_k\dot{F}_k(j\omega)\dot{S}_x(j\omega)+\sum_{k=0}^{2M-1}\dot{H}_k(j\omega)c_k\dot{F}_k(j\omega)\dot{N}(j\omega)\,, \qquad (1)$$

где $\dot{H}_k(j\omega)$, $F_k(j\omega)$ - комплексные частотные характеристики k-го фильтра анализа и k-го фильтра синтеза соответственно, $\dot{S}_x(j\omega)$ - спектральная плотность полезного сигнала, $\dot{N}(j\omega)$ - спектральная плотность реализации шума, c_k - коэффициент нормировки, значение которого обратно пропорционально значению средней мощности шума и помех в k-ом узкополосном канале.

Знак приблизительного равенства в (1) обусловлен отсутствием слагаемых, описывающих наложенные компоненты спектра, возникающие из-за снижения частоты дискретизации при обработке сигнала в каналах банка фильтров. Это при использовании фильтров с высоким уровнем подавления (свыше 100 дБ).

Примем, что согласованный фильтр (СФ) согласован только с полезным сигналом и учитывает только белый шум. Частотная характеристика согласованного фильтра равна

$$\dot{H}_{MF}(j\omega) = \dot{S}_x^*(j\omega)e^{-j\omega T_s} \tag{2}$$

Тогда спектральная плотность сигнала, прошедшего банк фильтров и СФ, равна

$$\dot{Y}_{MF}(j\omega) \approx \left|\dot{S}_x(j\omega)\right|^2 e^{-j\omega T_s}\sum_{k=0}^{2M-1}\dot{H}_k(j\omega)c_k\dot{F}_k(j\omega) + \dot{S}_x^*(j\omega)e^{-j\omega T_s}\sum_{k=0}^{2M-1}\dot{H}_k(j\omega)c_k\dot{F}_k(j\omega)\dot{N}(j\omega) \tag{3}$$

Отклик согласованного фильтра в момент снятия статистики может быть вычислен как

$$y_{MF}(t)\big|_{t=T_s+NT+0.5M(L-1)T} = \frac{1}{2\pi}\int_{-\infty}^{\infty}\dot{Y}_{MF}(j\omega)e^{j\omega\left(T_s+(N-1)T+0.5M(L-1)T\right)}d\omega, \tag{4}$$

где T_s - длительность обрабатываемого сигнала, $(N-1)T$ - задержка сигнала в банке фильтров, $0.5M(L-1)T$ - задержка сигнала в нормировщике в составе банка фильтров, T - интервал дискретизации широкополосного сигнала.

Отношение сигнал/шум на выходе СФ определяется как отношение квадрата модуля среднего значения отклика фильтра к его дисперсии

$$SNR = \frac{\left\langle y_{MF}(t)\big|_{t=T_s+NT+0.5M(L-1)T}\right\rangle^2}{D\left(y_{MF}(t)\big|_{t=T_s+NT+0.5M(L-1)T}\right)}. \tag{5}$$

Математическое ожидание и дисперсия отклика СФ для сигнала, прошедшего банк фильтров, равны соответственно

$$\left\langle y_{MF}(t)\big|_{t=T_s+NT}\right\rangle = \frac{1}{2\pi}\int_{-\infty}^{\infty}\left|\dot{S}_x(j\omega)\right|^2\sum_{k=0}^{2M-1}\dot{H}_k(j\omega)c_k\dot{F}_k(j\omega)e^{j\omega\left((N-1)T+0.5M(L-1)T\right)}d\omega, \tag{6}$$

$$D\left(y_{MF}(t)\big|_{t=T_s+NT}\right) = \frac{1}{2\pi}\int_{-\infty}^{\infty}G(\omega)\left|\dot{S}_x(j\omega)\right|^2\left|\sum_{k=0}^{2M-1}\dot{H}_k(j\omega)c_k\dot{F}_k(j\omega)\right|^2 d\omega, \tag{7}$$

где $G(\omega)$ - спектральная плотность мощности (СПМ) случайного процесса, описывающего мешающее воздействие, учитывающее шум и помехи.

Подставив (6), (7) в (5), получим выражение ОСШ на выходе СФ для сигнала, фильтрация узкополосных помех у которого, была осуществлена с помощью банка фильтров

$$SNR = \frac{\left(\int_{-\infty}^{\infty}\left|\dot{S}_x(j\omega)\right|^2\sum_{k=0}^{2M-1}\dot{H}_k(j\omega)c_k\dot{F}_k(j\omega)e^{j\omega\left((N-1)T+0.5M(L-1)T\right)}d\omega\right)^2}{\int_{-\infty}^{\infty}G(\omega)\left|\dot{S}_x(j\omega)\right|^2\left|\sum_{k=0}^{2M-1}\dot{H}_k(j\omega)c_k\dot{F}_k(j\omega)\right|^2 d\omega}. \tag{8}$$

Максимально достижимое ОСШ при обработке смеси сигнала с окрашенным гауссовским шумом определяется выражением

$$SNR = \frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{\left|\dot{S}_x(j\omega)\right|^2}{G(\omega)}d\omega. \tag{9}$$

Имея выражения (8) и (9) можно оценивать реально достижимый уровень ОСШ при использовании устройства подавления помех на базе банка цифровых фильтров и сравнивать с потенциально достижимым уровнем.

Результаты расчетов для натурных записей эфира

На рисунках 3 и 4 приведены результаты оценивания спектральной плотности мощности помех в полосе 160 кГц на различных частотах (14.6 МГц и 14.8 МГц в разное время) по записям эфира. На графиках представлены: спектральная плотность мощности помехи на входе банка цифровых фильтров, амплитудно-частотная характеристика банка цифровых фильтров и спектральная плотность мощности помехи на выходе банка цифровых фильтров.

АЧХ банка цифровых фильтров без учета наложения спектров рассчитана по формуле:

$$\left|\dot{H}_{БФ}(j\omega)\right| = \left|\sum_{k=0}^{2M-1} \dot{H}_k(j\omega) c_k \dot{F}_k(j\omega)\right| \qquad (10)$$

Коэффициенты c_k в (10) определялись путем вычисления средней мощности помехи в полосе пропускания соответствующих фильтров $\dot{H}_k(j\omega)$ и $\dot{F}_k(j\omega)$ по известной спектральной плотности мощности помехи $G(\omega)$.

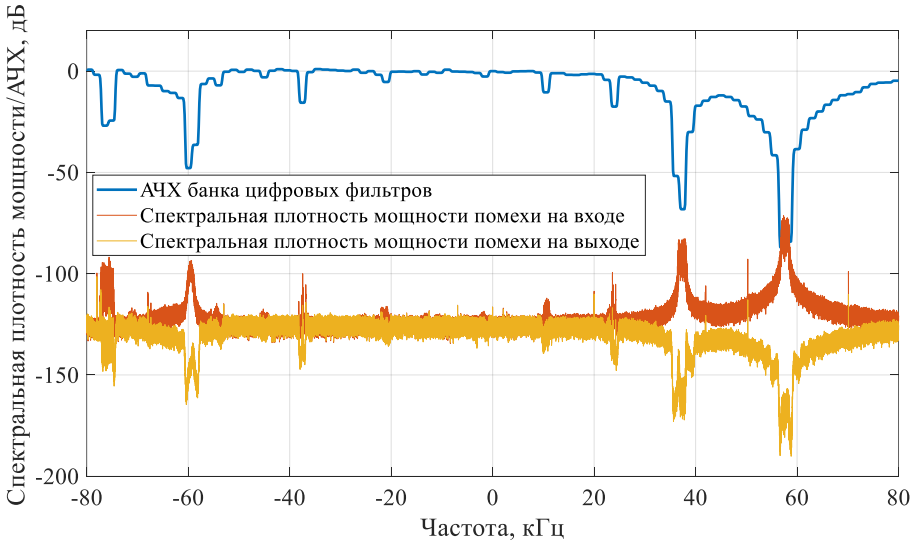


Рисунок 3 – СПМ помех на входе и выходе банка фильтров, АЧХ банка фильтров; 14.6 МГц

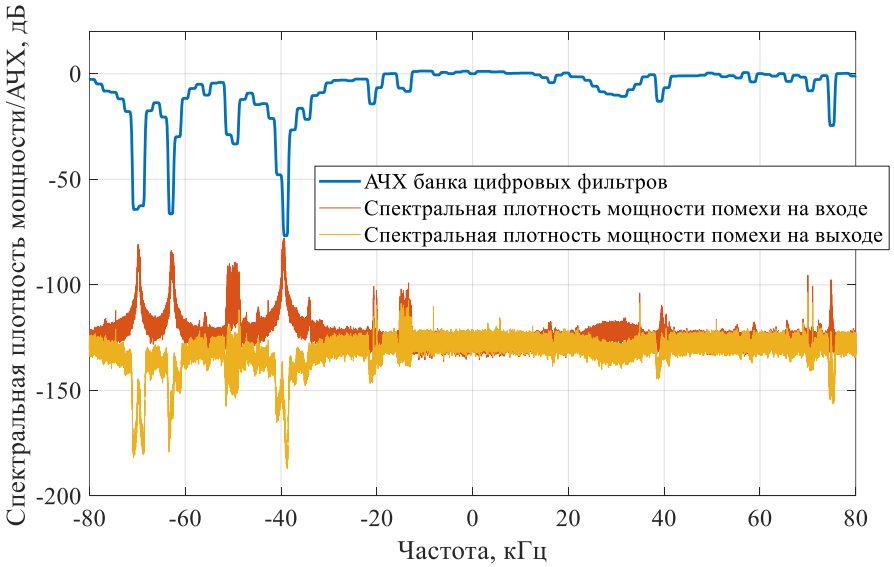


Рисунок 4 – СПМ помех на входе и выходе банка фильтров, АЧХ банка фильтров; 14.8 МГц

Обработка сигнала осуществлялась различным набором банком цифровых фильтров, состоящим из 2*M* фильтров от 16 до 1024 (степени двойки), которые разделяют основной диапазон частот, шириной 192 кГц (от -96 кГц до 96 кГц) на 2*M* узкополосных каналов.

Полезный принимаемый сигнал представляет собой широкополосный фазоманипулированный сигнал с символьной скоростью 100 кБод и шириной основного лепестка спектра 200 кГц, который при приеме фильтруется в полосе 160 кГц.

На рисунке 5 приведены кривые разности между потенциальным значением ОСШ и достижимым с помощью исследуемого устройства, рассчитанные с использованием формул (8) и (9), в зависимости от числа фильтров в банке цифровых фильтров 2M для исследуемых записей.

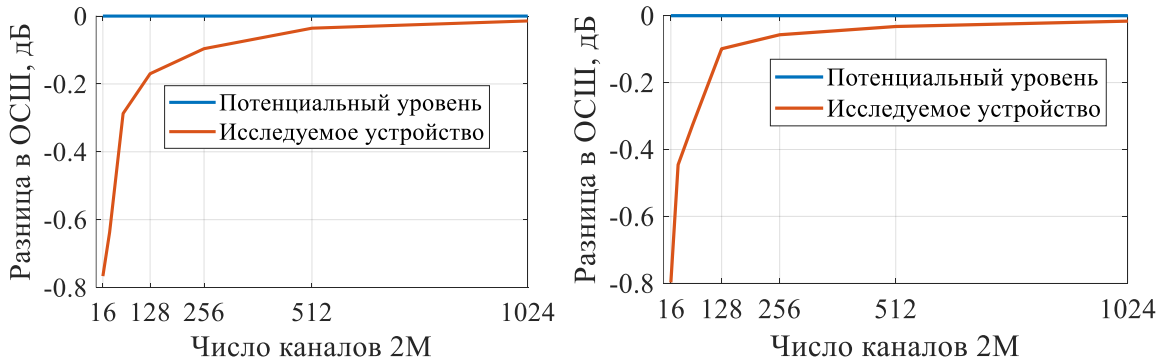


Рисунок 5 – Разница между потенциальным значением ОСШ и достижимым с помощью исследуемого устройства (слева для сигнала на частоте 14.6 МГц, справа – для 14.8 МГц)

Результаты расчетов показали, что разница между потенциально достижимым значением отношения сигнал/шум и достигаемом значением отношения сигнал/шум при использовании устройства подавления помех на базе банка цифровых фильтров составляет менее 0.2 дБ для обеих записей при использовании банка цифровых фильтров с числом каналов, равным 2M = 128, и менее 0.1 дБ при числе каналов 2M = 256. Указанная разность стремится к нулю при увеличении числа каналов в банке цифровых фильтров. Однако, вычислительная сложность устройства также увеличивается. В связи с изложенным рекомендуется использовать банк цифровых фильтров с числом каналов не более 128 на каждые 100 кГц полосы обрабатываемого сигнала.

Установленная разница объясняется отклонением АЧХ банка цифровых фильтров от АЧХ, точно равной обратной спектральной плотности мощности помех, которая уменьшается с увеличением числа каналов в банке цифровых фильтров.

Заключение

В работе представлены аналитические выражения, позволяющие рассчитать отношение сигнал/шум при квазиоптимальной обработке широкополосных сигналов с применением устройства подавления помех на базе банка цифровых фильтров.

Для натурных записей эфира расчетным путем показано, что разница между потенциально достижимым значением отношения сигнал/шум и достигаемом значением отношения сигнал/шум при использовании устройства подавления помех на базе банка цифровых фильтров составляет менее 0.2 дБ для обеих записей при использовании банка цифровых фильтров с числом каналов, равным 2M = 128, и менее 0.1 дБ при числе каналов 2M = 256.

Литература

1. Левин, Б. Р. Теоретические основы статистической радиотехники. Книга 2 / Б. Р. Левин. – М.: Советское радио, 1975. – 392 с.

2. Тихонов, В.И. Оптимальный прием сигналов / В.И. Тихонов. ¬¬¬– М.: Радио и связь, 1983. – 320 с.

3. Lobov E.M., Shubin D.N. A Narrow-Band Interference Compensation Device Based on a Digital Filter Bank for Broadband Low-Energy HF Radio Lines // 2019 Systems of Signals Generating and Processing in the Field of on Board Communications, Moscow, Russia, 2019, pp. 1-6, doi: 10.1109/SOSG.2019.8706791

4. A device for processing a wideband signal of the hf waveband using a digital filter bank / D.N. Shubin, E. M. Lobov, E. O. Lobova, N. A. Kandaurov // 2021 Systems of Signal Synchronization, Generating and Processing in Telecommunications, SYNCHROINFO 2021 - Conference Proceedings, Svetlogorsk, Kaliningrad Region, 30 июня – 02 2021 года. – Svetlogorsk, Kaliningrad Region, 2021. – P. 9488356. – DOI 10.1109/SYNCHROINFO51390.2021.9488356.

5. Свидетельство о государственной регистрации программы для ЭВМ № 2022615304 Российская Федерация. Программа режектирования помех на базе банка цифровых фильтров : № 2022614266 : заявл. 22.03.2022 : опубл. 30.03.2022 / Н. А. Кандауров, Д. Н. Шубин, Е. М. Лобов, В. И. Липаткин ; заявитель Ордена Трудового Красного Знамени федеральное государственное бюджетное образовательное учреждение высшего образования «Московский технический университет связи и информатики». – EDN YYZKJL.

6. В.В. Витязев, Многоскоростная обработка сигналов. М.: Горячая линия-Телеком, 2017, 336 с.

7. P.P. Vidyanathan, Multirate Systems and Filter Banks. Englewood Cliffs, NJ: Prentice-Hall, 1993.

8. Henrique S. Malvar. Signal Processing with Lapped Transforms. © 1992 ARTECH HOUSE, INC. ISBN 089006-467-9.

9. Lobov, E. M. Computationally simplified realization of the compensator of dispersion distortions on the basis of the filter bank / E. M. Lobov, E. O. Lobova, A. A. Kurochkin // 2018 Systems of Signal Synchronization, Generating and Processing in Telecommunications, SYNCHROINFO 2018, Minsk, 04–05 июля 2018 года. – Minsk: Institute of Electrical and Electronics Engineers Inc., 2018. – P. 8457058. – DOI 10.1109/SYNCHROINFO.2018.8457058

10. Lobov, E. M. Reducing the Volume of Broadband Signal Recordings Using a Digital Filter Bank / E. M. Lobov, D. N. Shubin, E. O. Lobova // 2021 23rd International Conference on Digital Signal Processing and its Applications, DSPA 2021 : 23, Moscow, 24–26 марта 2021 года. – Moscow, 2021. – DOI 10.1109/DSPA51283.2021.9535814.

11. Lobov, E. M. Investigation of the properties of an filterbank multicarrier modem in a broadband hf channel / E. M. Lobov, V. O. Varlamov // 2021 Systems of Signal Synchronization, Generating and Processing in Telecommunications, SYNCHROINFO 2021 - Conference Proceedings, Svetlogorsk, Kaliningrad Region, 30 июня – 02 2021 года. – Svetlogorsk, Kaliningrad Region, 2021. – P. 9488412. – DOI 10.1109/SYNCHROINFO51390.2021.9488412.

12. Офицерова, К. Б. Разработка и проектирование устройства предварительной фильтрации с использованием гибридных банков фильтров / К. Б. Офицерова, Е. М. Лобов // Телекоммуникации и информационные технологии. – 2019. – Т. 6, № 2. – С. 58-62. – EDN ADUARS.

13. Алгоритмы оценки и компенсации дисперсионных искажений широкополосных сигналов ионосферных радиолиний связи / С. С. Аджемов, Е. М. Лобов, Н. А. Кандауров [и др.] // Наукоемкие технологии в космических исследованиях Земли. – 2021. – Т. 13, № 5. – С. 57-74.

SIGNAL-TO-NOISE RATIO IN QUASI-OPTIMAL RECEPTION OF WIDEBAND SIGNALS USING A DIGITAL FILTER BANK BASED INTERFERENCE SUPPRESSION DEVICE

Lobov E.M., Shubin D.N., Grigorieva A.D.

Moscow technical university of communication and informatics

Expressions for calculating the signal-to-noise ratio for quasi-optimal reception of wideband signals using a noise suppression device based on a digital filter bank are obtained. Using broadcast recordings, calculations of the signal-to-noise ratio for signal reception in a 160 kHz band under narrowband interference conditions are performed. It is shown by calculation that the difference between the potentially achievable value of the signal-to-noise ratio and the achievable value of the signal-to-noise ratio when using a noise suppression device based on a digital filter bank with 256 filters is less than 0.1 dB for the processed recordings.

КОРРЕКЦИЯ ДИСПЕРСИОННЫХ ИСКАЖЕНИЙ В ИОНОСФЕРЕ: МЕТОДИКИ АНАЛИЗА И ПРАКТИЧЕСКОЕ ПРИМЕНЕНИЕ

асп. Галлямов К.Х., доц., доктор технических наук Лобов Е.М.

Московский технический университет связи и информатики

Ионосфера – это часть атмосферы Земли, находящаяся на высотах приблизительно от 30 до 1000 км, где происходит ионизация частиц под воздействием солнечного излучения. Ионосфера играет ключевую роль в радиосвязи, поскольку она отражает радиоволны, позволяя им преодолевать большие расстояния на сотни и тысячи километров. Как правило, используют широкополосные сигналы для более высокой скорости передачи информации, однако, различные факторы, включая дисперсионные искажения, могут значительно влиять на качество и точность радиосигналов. В этой статье рассматриваются методики, используемые для анализа и коррекции дисперсионных искажений в ионосфере.

Дисперсия сигналов в ионосфере

Осуществление радиосвязи по ионосферному каналу возможно в декаметровом диапазоне и довольно часто используется в задачах связи с труднодоступными регионами мира, например Арктикой или Антарктикой, военной сфере, связи с экстренными службами при чрезвычайных ситуациях [1,2]. В работе [3] обсуждается и потенциал внедрения такой связи для задач коммуникации с железнодорожной сетью. Немаловажная роль отводится широкополосным сигналам, поскольку это даёт возможность увеличения скорости передаваемой информации. Однако, наиболее острой проблемой здесь является неоднородность ионосферы, так как она является диспергирующей средой. Частотная дисперсия — это явление, при котором скорость распространения радиоволн зависит от их частоты. Если учесть и тот факт, что в зависимости от высоты изменяется её состав, в частности электронная концентрация [4], то это приводит к тому, что при прохождении через ионосферу сигналы разных частот могут приходить в разное время, вызывая искажения в передаваемой информации, как следствие, общая фазовая задержка носит нелинейный характер. Сказанное можно проиллюстрировать рисунком 1.

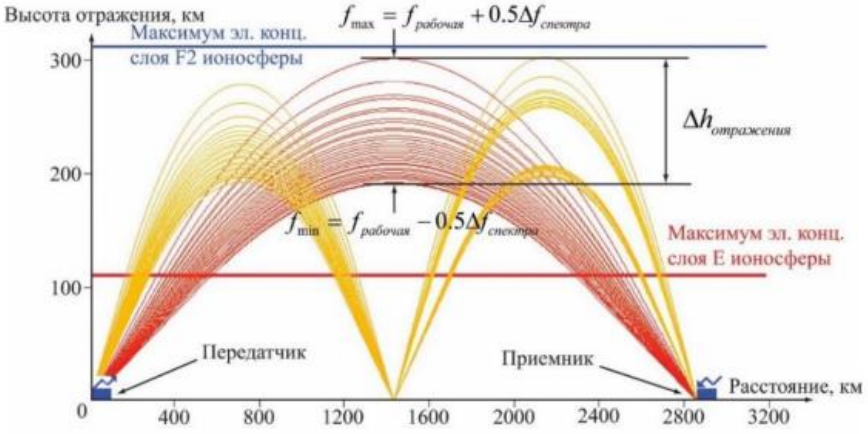


Рисунок 1 – Распространение радиоволн в ионосфере.

Как видно из рисунка 1, выделяют односкачковые и двухскачковые трассы. Как правило, односкачковые используют для расстояний, не превышающих 4000 км. Двухскачковые обычно используются в диапазоне от 4000 до 8000 км.

Модели описания дисперсионной характеристики

В следующем виде можно записать для i-го луча ионосферного канала частотную характеристику:

$$\dot{U}_i(f) = U_{i0}(f)e^{-j\varphi_i(f)}, \quad (1)$$

где $U_{i0}(f)$ – АЧХ для i-го луча, $\varphi_i(f)$ – ФЧХ для i-го луча.

Следует отметить, что введённое ранее понятие частотной дисперсии следует так же разделять на амплитудную дисперсию и фазовую дисперсию. Вторая, к слову, проявляется при нелинейности ФЧХ. Связь фазового сомножителя и времени распространения определяется по известному соотношению:

$$\tau_j(f) = \frac{1}{2\pi} d\varphi_j / df, \quad (2)$$

Зависимость, представленную в выражении (2) так же принято называть дисперсионной характеристикой (ДХ). На основе ДХ существует несколько моделей описания распространения сигналов в ионосфере [5]. Два из них представляют собой углубление в изучении радиофизики и нахождении приближенных решений уравнений геометрической оптики. Третьим же, интересующим нас вариантом будет полиномиальная модель, представленная в выражении (3).

$$\tau_j(f) = P_{nj}(f) = \sum_{k=1}^{n} \beta_k f^k \quad (3)$$

где $P_{nj}(f)$ – полином, β_k – коэффициент наклона касательной, n – максимальная степень полинома.

Она интересна простотой своего описания, а также опорой на натурные эксперименты, продуктами которых можно считать ионограммы, пример которой приведен на рисунке 2.

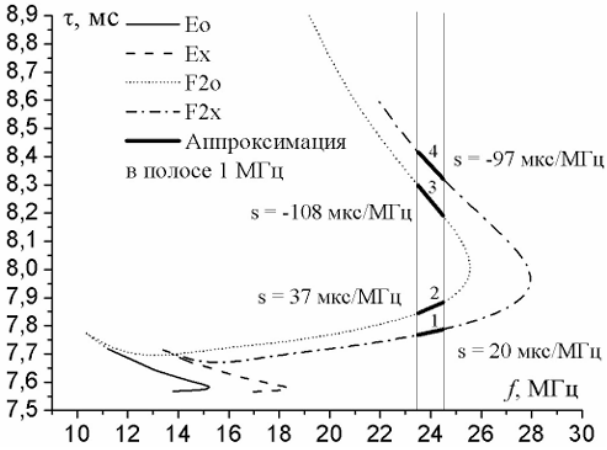


Рисунок 2 – Ионограмма Москва-Омск.

Обратившись к выражению (3), а также рисунку 1 следует отметить, что коэффициент β_k увеличивается в несколько раз ночью, а также при увеличении числа скачков при сравнении с односкачковой трассой. Степень полинома уменьшается с увеличением расстояния внутри одной моды. Как правило, максимальная степень полинома n ограничивается третьей степенью.

Перепишем выражение (3), выразив из него ФЧХ:

$$\varphi_j(f) = \varphi_j(f_\text{р}) + 2\pi \int_{f_\text{р}}^{f} \tau_{\text{н}j}(f)\, df, \quad (4)$$

где $f_\text{р}$ – рабочая частота, $\tau_{\text{н}j}(f)$ – групповая задержка наклонного распространения.

В случаях дальних трасс, а также вдали от максимально применимой частоты (МПЧ), если обратиться к рисунку 2, то это больше 2000 км и 30 МГц, можно ограничиться степенным рядом второй степени из выражения (4). Тогда ФЧХ можно переписать в виде:

$$\varphi_j(f) = \varphi_j(f_\text{р}) + 2\pi\tau_{0j}(f_\text{р})(f - f_\text{р}) + 2\pi s_j (f - f_\text{р})^2, \quad (5)$$

где s_j – крутизна ДХ.

Обратившись к работе [6], экспериментально полученные зависимости имеют параболическую зависимость в соответствии с рисунком 3 и рисунком 4.

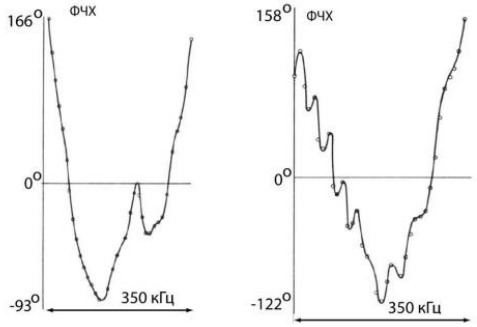


Рисунок 3 – Экспериментальная ФЧХ ионосферного канала в полосе 350 кГц: а) вдалеке от МПЧ, б) поблизости к МПЧ

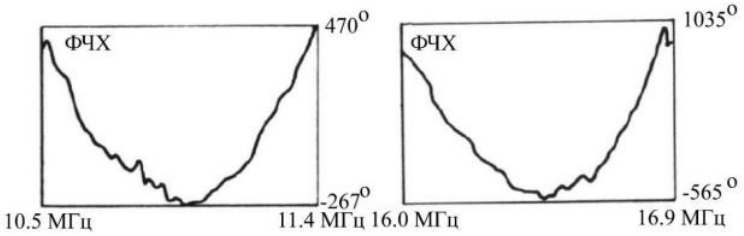


Рисунок 4 – Экспериментальная ФЧХ ионосферного канала в полосе 900 кГц: а) вдалеке от МПЧ, б) поблизости к МПЧ

Для того, чтобы определить искажение ионосферного канала вводят понятие когерентности сигнала, который в свою очередь является взаимосвязью между полосой широкополосного сигнала и набегом фазы равным 1 радиану относительно $f_{\text{р}}$. Её можно записать следующим образом:

$$\Delta f_{kj} = \sqrt{4/(\pi|s_j|)}, \tag{6}$$

Соответственно, по выражению (5) можно построить зависимость, представленную на рисунке 5. А по выражению (6) произвести расчёт полосы когерентности.

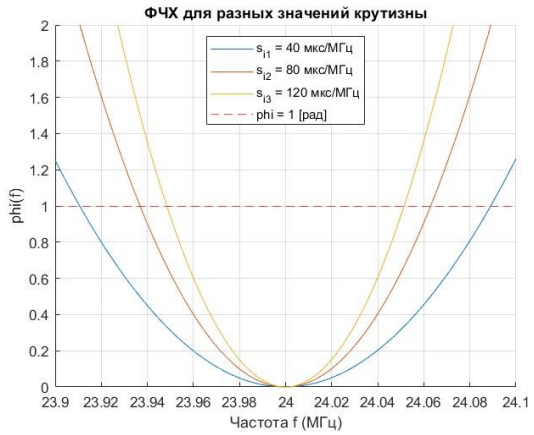


Рисунок 5 – ФЧХ при разной крутизне ДХ s_j.

Как можно заметить, с увеличением крутизны ДХ уменьшается полоса когерентности. Таким образом при крутизне 40мкс/МГц полоса когерентности составляет порядка 178кГц, а при 120 мкс/МГц всего около 100кГц. Критическое значением крутизны ДХ можно представить в соответсвии с выражением (7):

$$s_k = 4/(\pi\Delta f^2), \tag{7}$$

где Δf – полоса пропускания.

Тогда коэффициент дисперсности можно определить таким образом:

$$p_j = \Delta f/\Delta f_{kj}, \tag{8}$$

Так же стоит обратить внимание на зависимость среднего наклона ДХ от длины трассы. Данная зависимость проиллюстрирована на рисунке 6.

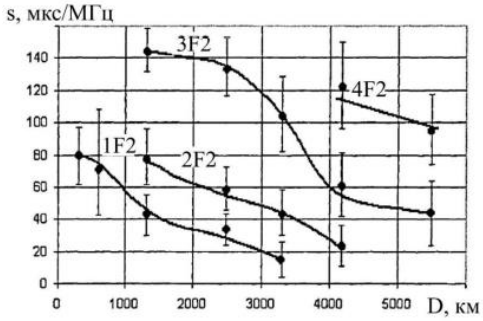


Рисунок 6 – Зависимость средней крутизны ДХ от дальности.

Как можно заметить, с увеличением дальности трассы в рамках одной моды, средняя крутизна ДХ уменьшается. Другой особенностью является оценка потерь в решении задачи обнаружения

детерминированного комплексного сигнала с начальной фазой. Если на выходе СФ ОСШ можно представить в виде:

$$\frac{P_{\text{с}}}{P_{\text{ш}}} = \frac{E_{\text{с}}}{N_0} K^2(s), \tag{9}$$

где $K(s)$- коэффициент потерь.

График зависимости потерь от наклона ДХ представлен на рисунке 7.

В соответствии, а также построив график зависимости коэффициента потерь от крутизны ДХ, представлена на рисунке 7.

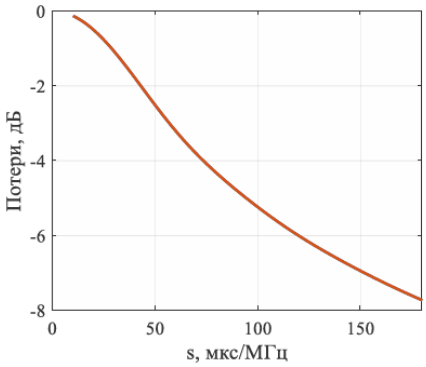


Рисунок 7 – Зависимость $K(s)$ от крутизны ДХ.

Анализируя выражение (9), а также зависимость, реализованную на рисунке 7, можно отметить уменьшение ОСШ, что сказывается негативным образом на качестве обнаружения сигнала и приеме информации при увеличении крутизны ДХ.

Способы коррекции дисперсионных искажений

Как было сказано ранее, частотную дисперсию принято разделять на амплитудную и фазовую. Наиболее распространенными алгоритмами компенсации искажений, вызванных в контексте амплитудной частотной дисперсии, являются: обратная фильтрация, согласованная фильтрация и фильтр Винера [7]. Однако, оптимальными являются лишь последние два.

Для обратной фильтрации передаточная функция имеет вид:

$$K_{\text{оф}}(j2\pi f) = \frac{1}{U(j2\pi f)}, \tag{10}$$

Данный фильтр нельзя назвать оптимальным, поскольку не принимаются во внимание спектральные характеристики сигнала, распространяющегося по каналу, а также изменения спектральной плотности помех.

При рассмотрении компенсатора на основе согласованной фильтрации необходимо согласовывать фильтр и с полезным сигналом, и с каналом, тогда передаточную функцию можно записать в виде:

$$K_{\text{СФ}}(j2\pi f) = \frac{X^*(j2\pi f)U^*(j2\pi f)}{N(f)}, \tag{11}$$

где $N(f)$- спектральная плотность мощности шума, $X(j2\pi f)$ – спектр передаваемого сигнала. Алгоритм согласованной фильтрации оптимален по критерию максимума ОСШ.

Фильтр Винера является восстановительным статистическим алгоритмом. Его задачей является минимизация среднеквадратичной ошибки (СКО), и оптимален по данному критерию. Как правило, используется в случаях с узкополосными провалами в спектре на основе анализа информации поведения сигнала на других частотах. Передаточная функция такого фильтра записывается в виде:

$$K_{\text{В}}(j2\pi f) = \frac{U(j2\pi f)}{|U(j2\pi f)|^2 + TN(f)/|X(j2\pi f)|^2}, \tag{12}$$

где T – длительность сигнала.

Представленные выше алгоритмы пригодны для случаев амплитудной частотной дисперсии. Однако, в работе [8] предложен компенсатор, находящийся в составе банка цифровых фильтров, связывающий отклонение ФЧХ по двум оптимальным критериям: среднеквадратический и минимаксный. Структурная схема такого банка цифровых фильтров со встроенным алгоритмом компенсации представлена на рисунке 8.

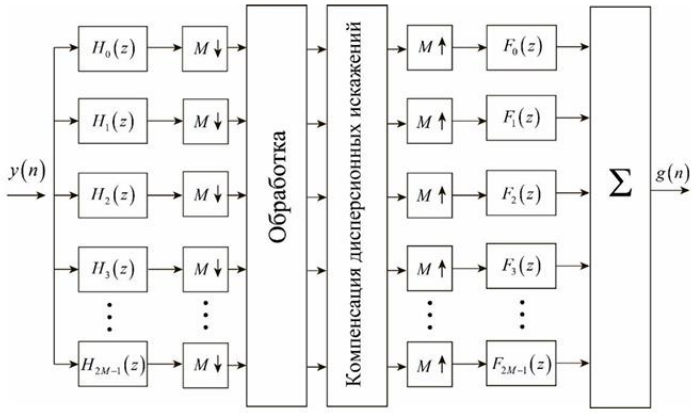


Рисунок 8 – Структурная схема банка фильтров со встроенным алгоритмом компенсации дисперсионных искажений.

Такой банк включает в себя 2*M* фильтров анализа, 2*M* дециматоров, блок обработки на низкой частоте, алгоритм компенсации дисперсионных искажений, 2*M* интерполяторов, 2*M* фильтров-синтеза и сумматор.

Одним из условий такого компенсатора является сохранение равномерности АЧХ в тракте по возможности при компенсации нелинейной составляющей ФЧХ. Из результатов работы [8] можно отметить, что реализация банка фильтров с внедренным алгоритмом компенсации является более вычислительно эффективной, чем реализация банка фильтров и фильтра-компенсатора как два отдельных устройства. При одинаковых требованиях к выравниванию ФЧХ, по среднеквадратическому критерию $\varphi_{\text{скв}}$ математических операций требуется в 1,5 раза меньше по сравнению с фильтром-компенсатором, по чебышевскому (минимаксному) $\varphi_{\text{чеб}}$ в 2 раза меньше соответственно.

Первый случай был рассмотрен для полного синтеза, то есть все каналы задействованы. В случае, когда нас интересует известный нам узкий диапазон, тогда имеется возможность некоторые каналы не задействовать. Структурная схема такого компенсатора представлена на рисунке 9.

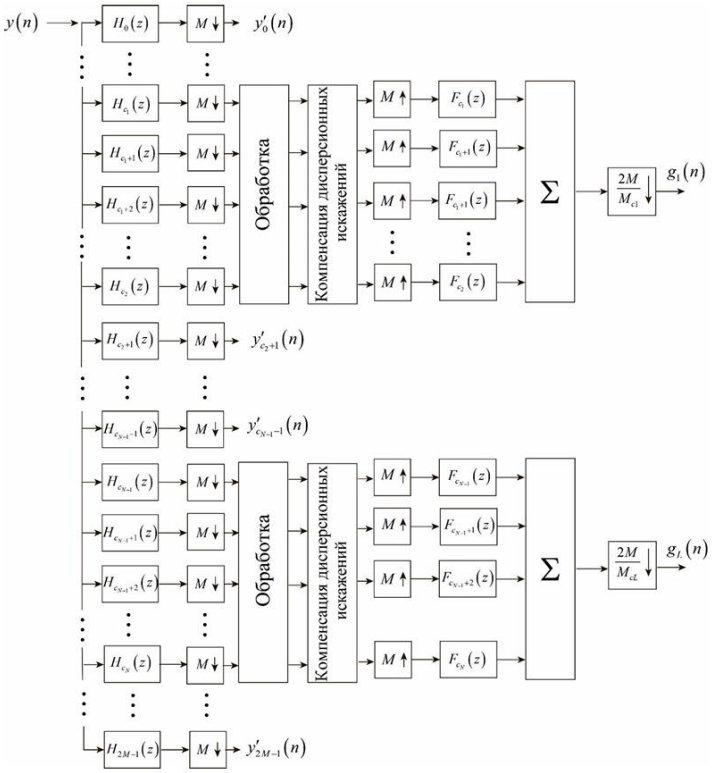


Рисунок 9 – Структурная схема банка фильтров со встроенным алгоритмом компенсации дисперсионных искажений для случая частичного синтеза.

Как следует из работы [8] число математических операций одинаково для полного и частичного синтеза, как для банка цифровых фильтров со встроенным алгоритмом компенсации, так и для случая, когда банк цифровых фильтров и фильтр-компенсатор являются разными устройствами. Следовательно, выводы, сделанные для полного синтеза так же справедливы и для частичного.

Литература

1. Kwasinski, A. Telecommunications power plant damage assessmentcaused by Hurricane Katrina – Site survey and follow-up results / A. Kwasinski, W.W. Weaver, P.L. Chapman, P.T. Krein // INTELEC. Providence, RI, USA. – 2006. – 8 p.
2. Коваль С. А. Ионосферный мониторинг в интересах перспективных адаптивных систем декаметровой радиосвязи: современное состояние и перспективы развития // Системы управления, связи и безопасности. – 2020 – №. 4 – С. 73-100.
3. Лапунов С.И., Роенков Д.Н., Плеханов П.А., Глухов И.А. Применение систем коротковолновой радиосвязи на малоинтенсивных железнодорожных участках // Автоматика, связь, информатика. 2022. № 9. С. 4–8. DOI: 10.34649/AT.2022.9.9.001
4. Гололобов, Д. В. Распространение радиоволн и антенно-фидерные устройства: метод. пособие для студ. спец. 45 01 02 «Системы радиосвязи, радиовещания и телевидения» дневной и вечерней форм обучения: В 3 ч. Ч. 1. Распространение радиоволн / Д. В. Гололобов, В. Б. Кирильчук. – Минск : БГУИР, 2003. – 124 с.: ил.
5. Иванов, Д.В. Развитие радиофизических методов и математических моделей исследования распространения в ионосфере сложных декаметровых сигналов и коррекции их дисперсионных искажений: диссертация ... доктора физико-математических наук: 01.04.03 / Иванов Дмитрий Владимирович; [Место защиты: Ин-т радиотехники и электроники РАН]. - Йошкар-Ола, 2007. - 329 с.: ил.
6. Perry, B.D. Equalized Megahertz-Bandwidth HF Channels for Spread Spectrum Communications / B.D. Perry, S. Dhar // IEEE Military Communications Conference. – 1982. – P. 29.5-1 - 29.5-5.
7. Исследование дисперсности сверхширокополосных декаметровых радиоканалов и коррекции дисперсионных искажений / Иванов В.А., Иванов Д.В., Лащевский А.Р. II III Всерос. науч. конф. «Сверхширокополосные сигналы в радиолокации связи и акустике». - Муром. - 2010 -С.7-16.
8. Лобова, Е.О. Алгоритмы компенсации дисперсионных искажений широкополосных сигналов на базе банка цифровых фильтров: диссертация ... кандидата технических наук: 05.12.04 / Лобова Елизавета Олеговна; [Место защиты: ОТКЗ ФГБОУ ВО «Московский технический университет связи и информатики»]. — Москва, 2021. — 189 с.: ил.

CORRECTION OF DISPERSION DISTORSIONS IN THE IONOSPHERE: METHODS OF ANALYSIS AND PRACTICAL APPLICATION

Gallyamov K.H., Lobov E.M.

Moscow Technical University of Communications and Informatics

The ionosphere is a part of the Earth's atmosphere located at altitudes of approximately 30 to 1,000 km, where particles are ionized by solar radiation. The ionosphere plays a key role in radio communications, as it reflects radio waves, allowing them to travel long distances of hundreds and thousands of kilometers. As a rule, broadband signals are used for higher information transmission rates, however, various factors, including dispersion distortions, can significantly affect the quality and accuracy of radio signals. This article discusses techniques used to analyze and correct dispersion distortions in the ionosphere.

ГРАНИЦЫ ИССЛЕДОВАНИЯ ПРОЦЕССА НАСТРОЙКИ РЕГУЛИРУЕМЫХ ПАРАМЕТРОВ ПРОТОКОЛА ПЕРЕДАЧИ ДАННЫХ ПОДУРОВНЕМ УПРАВЛЕНИЯ ТРАФИКА В ПЕРСПЕКТИВНОМ ПАКЕТНОМ ДЕКАМЕТРОВОМ РАДИОКАНАЛЕ

д.т.н., доц. Попов М.Ю.[1], к.т.н. Слободсков О.Е.[2],
асп. Киреев Д.А,[1] асп. Торбенко Я.К.[1]

[1]Филиал военной академии РВСН имени Петра Великого в г. Серпухове.
[2]АО «Российский институт мощного радиостроения» (г. Санкт-Петербург).

В статье представлена система допущений, ограничений и правил, устанавливающих границы исследования процесса настройки регулируемых параметров протокола передачи данных подуровнем управления трафика в пакетном декаметровом радиоканале общего назначения. Показана детализированная схема управления информационным обменом внутренним и внешним подпроцессом.

Для дальнейшей формализации процесса настройки регулируемых параметров (РП) протокола передачи данных (ППД) и построения математической модели введем следующие необходимые для этого определения, допущения и ограничения, то есть установим границы предстоящего исследования [1-3].

Определение 1. **Канал обработки первого типа** – канал обработки сообщения, реализующий первый тип обслуживания.

Определение 2. **Канал обработки второго типа** – канал обработки сообщения, реализующий второй тип обслуживания.

Определение 3. **Сообщение** – сообщение, содержащее значения регулируемых параметров ППД, определенных в ходе оценки качества канала радиосвязи (количество пакетов и количество переносящих ими бит информации, количество переспросов непринятых пакетов), предназначено для установления РП ППД на сеансе связи.

Определение 4. **Первый тип обслуживания** – процесс, реализующий предварительную настройку регулируемых параметров ППД.

Определение 5. **Второй тип обслуживания** – процесс, реализующий непосредственную настройку ППД, согласно полученных значений РП.

Ограничение 1. Канал обработки (КО) первого типа имеет 1 канал обслуживания, канал обработки второго типа имеет 2 канала обслуживания.

Ограничение 2. Буфер КО 1-го типа имеет емкость 1, буфер КО 2-го типа имеет емкость 2.

Допущение 1. Предварительная настройка РП ППД – процесс осуществления контроля наличия конфликта устанавливаемых регулируемых параметров ППД - прерывается в случае его выявления.

Допущение 2. Непосредственная настройка РП ППД – процесс установления параметров ППД на сеансе связи в виде количества пакетов, их объема и количества переспросов непринятых пакетов - может прерваться в случае наличия признака внешнего прерывания.

Допущение 3. Сообщения перед поступлением в каналы обработки всегда поступают в соответствующие буферы.

Допущение 4. Сообщение, поступающие в канал обработки второго типа, может обрабатываться в любом одном из свободных каналов обслуживания.

Допущение 5. Каналы обслуживания каналов обработки могут использоваться другими процессами, в следствии чего они могут быть недоступны для обработки сообщения.

Допущение 6. Сообщение в буферах может находиться до момента высвобождения каналов обслуживания.

Ограничение 3. Процесс обработки сообщения в КО 1 типа прерывается в случае наличия признака конфликта РП.

Ограничение 4. Процесс обработки сообщений в КО 2 типа прерывается в связи с наличием признака внешнего прерывания процесса.

Определение 6. Конфликт РП – выявленное (процессом более высокого уровня) противоречие настраиваемых значений РП ППД с текущими настройками процесса информационного обмена.

Следствие наличия конфликта РП – невозможность дальнейшего использования найденных РП в ППД.

Определение 7. Внешнее прерывание – команда, от процесса более высокого уровня, на отмену настройки РП ППД.

Ограничение 5. Внешнее прерывание возникает в следствии наличия:

- команды «стоп» оператора автоматизированного рабочего места (АРМ) автоматизированной системы управления (АСУ) (инициализируется внутренним процессом);
- перерыва радиосвязи 2 и 3 типа при условии появления их в момент осуществления фазы настройки ППД (инициализируется внешним процессом).

Необходимо также отметить то, что при обслуживании сообщения в КО 1 и 2 типов может возникнуть необходимость их использования для обслуживания более приоритетных сообщений, поэтому имеется возможность выгрузки из КО служебного сообщения, содержащего РП ППД в соответствующие буфера, что влечет за собой появления задержки обслуживания служебного сообщения.

Определение 8. Задержка обслуживания служебного сообщения – процесс временной выгрузки служебного сообщения из КО, с целью их высвобождения для осуществления обработки более приоритетных сообщений.

Задержка обслуживания служебного сообщения характеризуется соответствующим временем.

В связи с изложенным введем следующие ограничение.

Ограничение 6. Задержка обслуживания служебного сообщения возникает в следствии наличия необходимости (инициализируется внутренним процессом) предоставления КО для обработки сообщения, имеющих более высокий приоритет обслуживания.

Необходимо также отметить, что ухудшение помеховой обстановки при осуществлении информационного обмена в декаметровом радиоканале, возможно, как в следствии применения злоумышленником радиоэлектронного подавления, так и в силу естественных особенностей декаметровой радиосвязи [4,5] (наличие медленных и быстрых замираний и т.п.).

Одним из параллельных подпроцессов работы декаметрового радиоканала является процесс принудительного отключения приемного устройства и смежных систем (в частности аппаратуры передачи данных) в случае наличия электромагнитного импульса большой мощности, природа возникновения которого может быть различна (как естественная, например, гроза, так и искусственная).

Данные факторы влияют на качество информационного обмена в радиоканале (от ухудшения, до полного прекращение приема информации в нем), следовательно, текущие, определенные на сеансе связи регулируемые параметры ППД могут стать не актуальными (не обеспечивать требуемые вероятностно-временные характеристики доведения информации) и возможно, потребуется:

- их переопределение на текущей рабочей частоте,
- уход с данной рабочей частоты и поиск другой, пригодной для работы, с последующей настройкой РП ППД.

Таким образом, ухудшения помеховой обстановки, в силу приведенных естественных и искусственных факторов, может вызывать перерывы в радиосвязи и в зависимости от значения времени этого перерыва ($t_{прс}$) существуют несколько альтернатив продолжения процесса информационного обмена в радиоканале, которые возможно описать в виде следующего правила.

Правило 1. Если $t_{прс}$:

1. $0 < t_{прс} < t_{прс}^{и}$ - осуществляется возобновление информационного обмена на имеющейся рабочей частоте (РЧ) и текущих регулируемых параметров ППД;
2. $t_{прс}^{и} \le t_{прс} < t_{прс}^{с}$ - осуществляется возврат к фазе настройки РП ППД на текущей РЧ;

3. $t^{c}_{прс} \le t_{прс} < t^{в}_{прс}$ - осуществляется возврат к фазе анализа качества радиоканала - поиск оптимальной РЧ,

где $t^{н}_{прс}$ - нижняя граница времени перерыва радиосвязи, [с.];

$t^{c}_{прс}$ - средняя граница времени перерыва радиосвязи, [с.];

$t^{в}_{прс}$ - верхняя граница времени перерыва радиосвязи, [с.].

Для дальнейшей формализации процесса введем следующие определения.

Определение 9. Перерыв радиосвязи 1 типа – перерыв в радиосвязи, при котором $t_{прс} \in \left(0; t^{н}_{прс}\right)$.

Определение 10. Перерыв радиосвязи 2 типа – перерыв в радиосвязи, при котором $t_{прс} \in \left[t^{н}_{прс}; t^{c}_{прс}\right)$.

Определение 11. Перерыв радиосвязи 3 типа – перерыв в радиосвязи, при котором $t_{прс} \in \left[t^{c}_{прс}; t^{в}_{прс}\right)$.

В связи с изложенным, детализированная схема управления информационным обменом по пакетному радиоканалу внутренним и внешним процессом, в силу приведенных определений, допущений и ограничений, будет выглядеть так - рисунок 1.

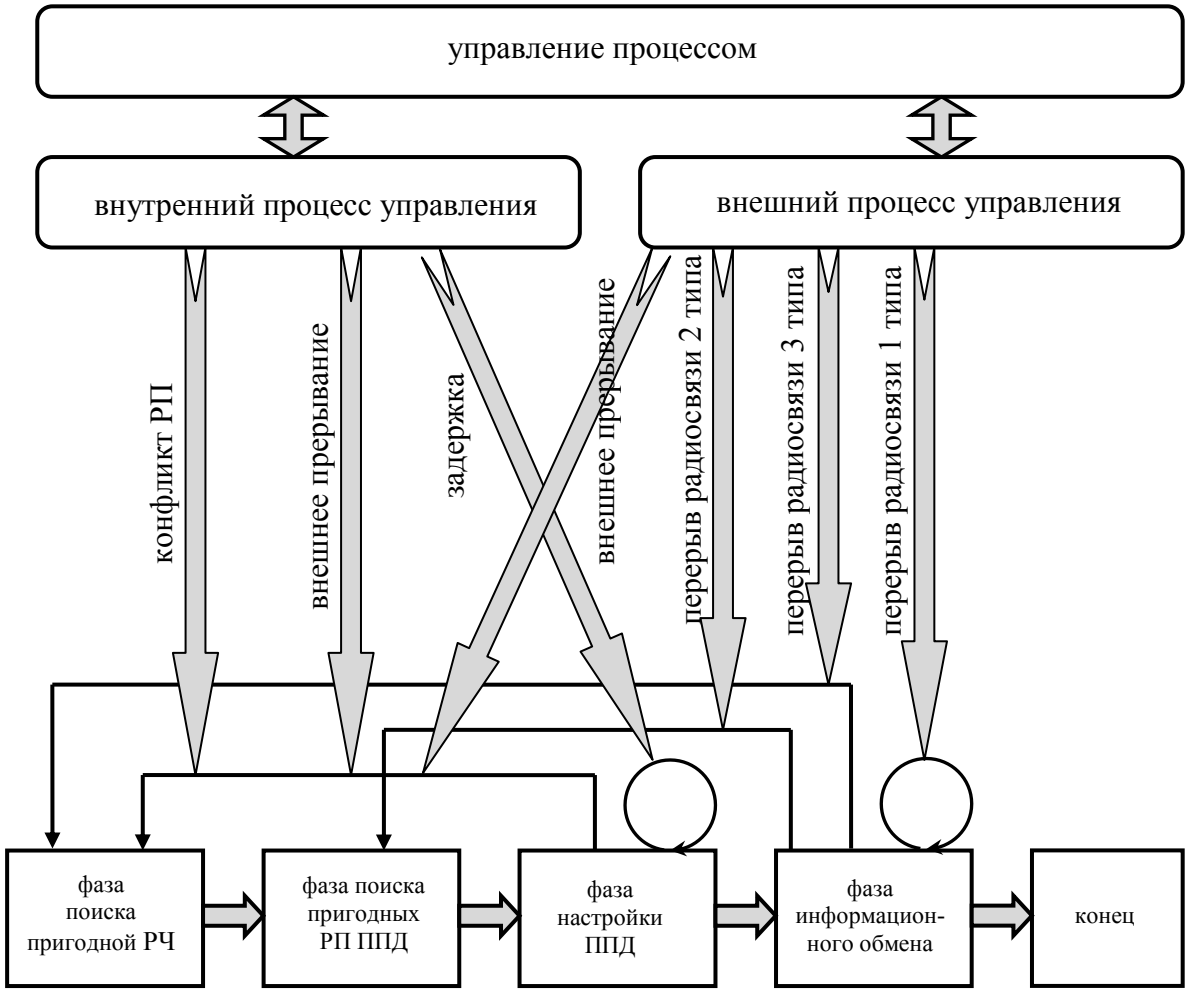


Рисунок 1 - Детализированная схема управления информационным обменом внутренним и внешним процессом

Таким образом, приведены границы исследования процесса настройки регулируемых параметров протокола передачи данных подуровнем управления трафика в перспективном пакетном декаметровом радиоканале общего назначения с соответствующей схемой управления информационным обменом.

Литература

1. Формализация процесса обслуживания абонента системой обеспечения вызова экстренных оперативных служб по единому номеру «112» конечной полумарковской цепью [Текст] / В. А. Цимбал, М. Ю. Попов, Д. А. Киреев, Н. В. Попова, И. Б. Иванов // Проектирование и технология электронных средств : всерос. науч.–техн. жур. – Владимир, 2023. – № 2. – С. 12–20. – ISSN 2071–9809.

2. Обоснование регулируемых параметров модели процесса обслуживания абонента системой обеспечения вызова экстренных оперативных служб по единому номеру «112» [Текст] / М. Ю. Попов, В. Б. Девятияров, Д. А. Киреев, Н. В. Попова, И. Б. Иванов // Проектирование и технология электронных средств : всерос. науч.–техн. жур. – Владимир, 2023. – № 2. – С. 52–57. – ISSN 2071–9809.

3. Описание регулируемых параметров модели, созданной на базе полумарковской цепи, описывающей действия оператора системы обеспечения вызова экстренных оперативных служб по единому номеру «112» [Текст] / М. Ю. Попов, С. С. Чайков, Н. В. Попова, С. А. Кравец, И. Б. Иванов // Доклады 25-й международной конференции «Цифровая обработка сигналов и её применение» (DSPA-2023); Серия: цифровая обработка сигналов и ее применение / Рос. науч.-техн. общ. радиотехн., электрон. и связи им. А.С. Попова. – М., 2023. - Вып. XXV.,– С. 51–55. – ISBN 978-5-905278-53-2.

4. Регулярные конечные марковские цепи, описывающие процесс доведения сообщения объекту управления критической инфраструктуры по радиоканалу управления автоматизированной системы управления в условиях энергетического и информационного воздействия [Текст] / М. Ю. Попов, Е. В. Калганов, Н. В. Попова, Д. А. Киреев, И. Б. Иванов // Проектирование и технология электронных средств : всерос. науч.–техн. жур. – Владимир, 2023. – № 1. – С. 16–23. – ISSN 2071–9809.

5. Математическая модель процесса доведения сообщения по радиоканалу управления специальной АСУ в условиях энергетического и информационного воздействия на базе конечной марковской цепи с доходами [Текст] / В. А. Цимбал, М. Ю. Попов, В. В. Илющенко, Е. В. Калганов, А. М. Деркач // Доклады 24-й международной конференции «Цифровая обработка сигналов и её применение» (DSPA-2022); Серия: цифровая обработка сигналов и ее применение / Рос. науч.-техн. общ. радиотехн., электрон. и связи им. А.С. Попова. – М. : Вып. XXIV.,– С. 70–75. – ISBN 978-5-905278-49-5.

BORDERS OF RESEARCH OF PROCESS OF ADJUSTMENT OF ADJUSTABLE PARAMETERS OF THE REPORT OF DATA TRANSMISSION BY THE SUBTOTAL OF MANAGEMENT OF THE TRAFFIC IN PERSPECTIVE PACKAGE SHORT-WAVE THE RADIO CHANNEL

Dr.Sci.Tech. Popov M.Yu.[1], Cand.Tech.Sci. Slobodskov O.E.[2],
Kireev D.A.[1], Torbenko Ya.K.[1]

[1]Branch of military academy of Rocket strategic forces of a name of Peter the Great
(c. Serpukhov, Moscow region).
[2]The Russian institute of a powerful radio structure, St.-Petersburg.

In article the system of assumptions, restrictions and the rules establishing borders of research of process of adjustment of adjustable parametres of the report of data transmission by a subtotal of management of the traffic in a package short-wave radio channel of a general purpose is presented. The detailed scheme of management is shown by an information exchange of internal and external process.

ФОРМАЛИЗАЦИЯ ПРОЦЕССА УСТАНОВЛЕНИЯ СОЕДИНЕНИЯ И ОСУЩЕСТВЛЕНИЯ ИНФОРМАЦИОННОГО ОБМЕНА В ПАКЕТНОМ РАДИОКАНАЛЕ НА БАЗЕ TCP-ПОДОБНОГО СОЕДИНЕНИЯ

д.т.н., доц. Попов М.Ю.[1], к.т.н., доц. Кабанович С.Г.[1],
асп. Торбенко Я.К.[1], Попов Р.В.[2]

[1]Филиал военной академии РВСН имени Петра Великого, г. Серпухов
[2]16 ЦНИИИ МО РФ, г. Москва

В статье на базе конечного автомата, временных диаграмм и диаграммы Хассе представлена формализация процесса установления соединения и осуществления информационного обмена в пакетном радиоканале на базе TCP-подобного соединения. Описан физический смысл соответствующих состояний и переходов. Дана система необходимых определений, ограничений и допущений.

При рассмотрении задач, направленных на оценку и повышение качества информационного обмена в радиоканалах и сетях передачи данных, всегда необходимо учитывать соответствующие фазы процесса, такие как: установление соединения, информационный обмен и разрыв соединения. Данные фазы содержат процедуры и правила, описывающие их поведение. С целью решения обозначенных вопросов необходимо отталкиваться от существующих правил установления соединения. Наиболее рельефно они изложены в протоколе TCP, применяемом в качестве транспортного протокола во множестве приложений Internet. Он обеспечивает надежное отслеживание потоков байт в сегментах, каждый из которых передается как дейтаграмма IP (Internet Protocol) [1].

Известна следующая диаграмма состояний TCP-соединения, представленная в виде конечного автомата, описывающего процесс информационного обмена в TCP соединении (рисунок 1) [1].

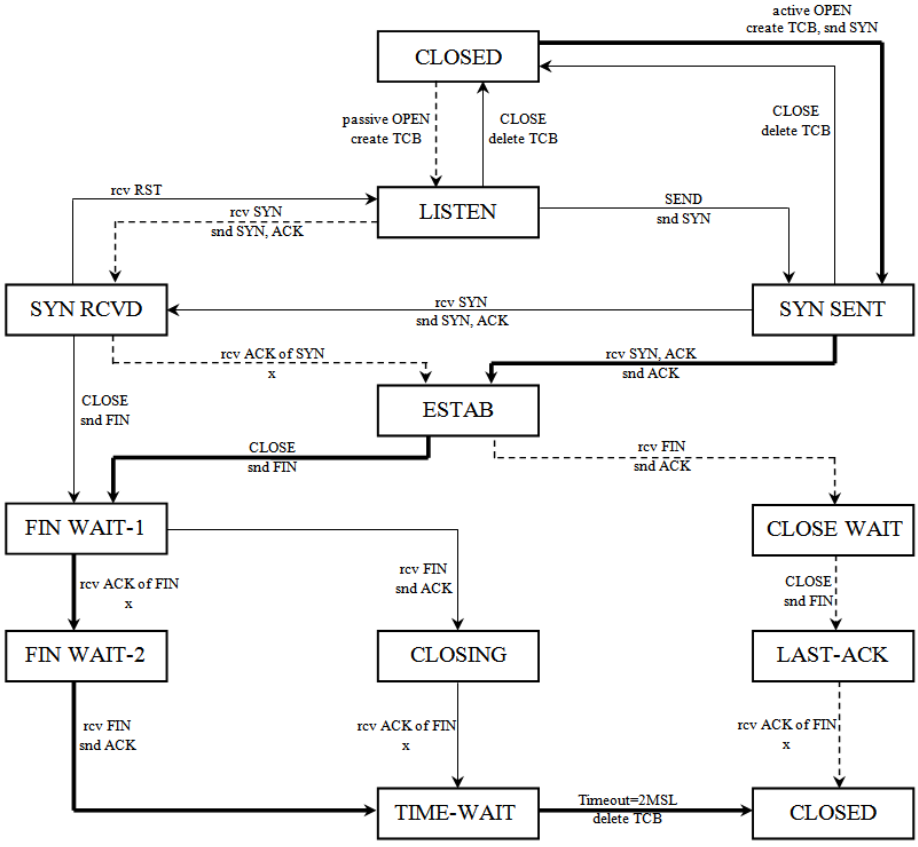


Рисунок 1 – Конечный автомат процесса информационного обмена в TCP соединении

Частный случай организации информационного обмена (3-этапное согласование для установления соединения, а также случай нормального закрытия при завершении соединения) представлен в виде временной диаграммы на рисунке 2.

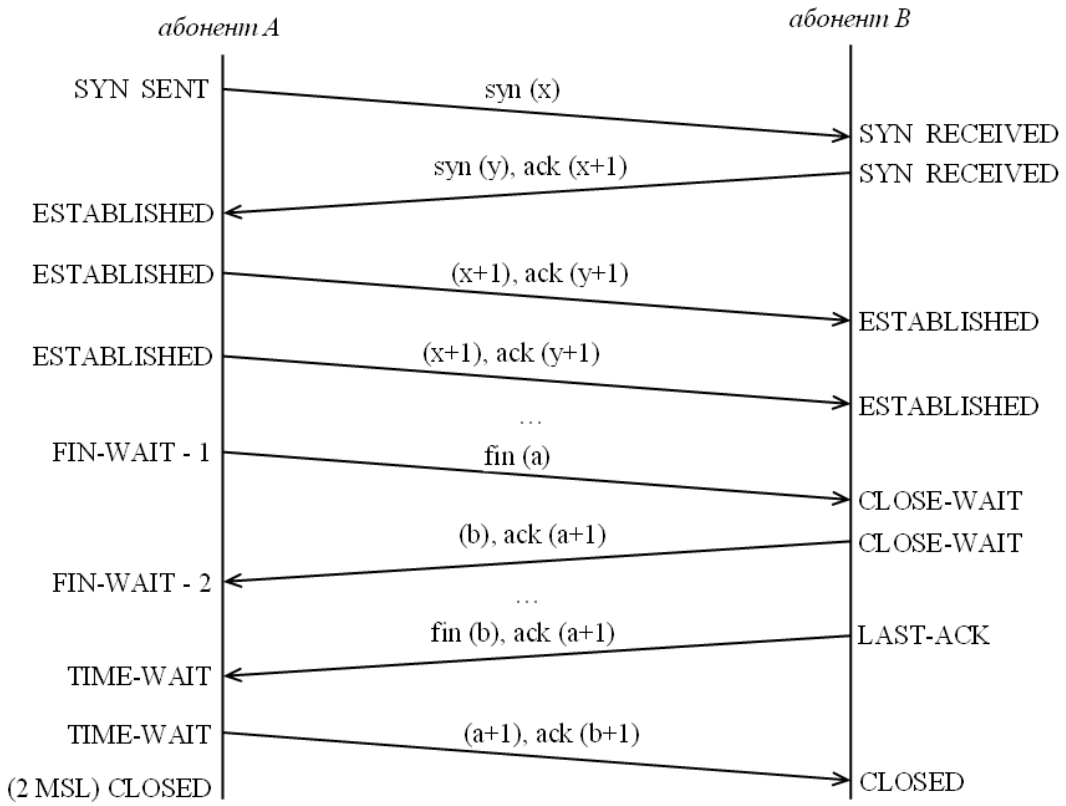


Рисунок 2 – Временная диаграмма процесса информационного обмена в TCP соединении

Физический смысл состояний автомата таков:

1. **CLOSED** – состояние, соответствующее отсутствию соединения или завершению времени таймаута после окончания предыдущего соединения и удаленной записи соединения из соответствующего буфера или абонентом B получено подтверждение от абонента A на завершение существующего соединения (абоненту B нечего передавать);
2. **LISTEN** – состояние, соответствующее осуществлению поиска сигналов сообщений абонентами;
3. **SYN SENT** – состояние, соответствующее ожиданию подтверждения от абонента B на установление соединения с абонентом A (запрос от абонента A отправлен);
4. **SYN RCVD** – состояние, соответствующее ожиданию подтверждения на установление соединения от абонента A (получен запрос от абонента A, передан запрос от абонента B и подтверждение для абонента A) (сразу после отправки запроса абонентом A, пришел запрос от абонента B. Абонент B подтвердил запрос абонента A и ожидает подтверждения своего запроса);
5. **ESTAB** – состояние, соответствующее информационному обмену между абонентами;
6. **FIN WAIT-1** – состояние, соответствующее инициации закрытия соединения (абонент A отправил запрос на закрытие соединения);
7. **FIN WAIT-2** – состояние, соответствующее прекращению передачи информации абонентом A (получено подтверждение запроса на разрыв соединения от абонента B);
8. **CLOSING** – состояние, соответствующее ожиданию подтверждения от абонента B на закрытие соединения (абонент A получил запрос на завершение соединения от абонента B до

получения подтверждения на свой запрос. Абонент А отправил абоненту В подтверждение на завершение соединения);

9. **TIME-WAIT** – состояние, соответствующее завершению процедуры разрыва соединения (выжидание таймаута на случай запаздывания сегментов);

10. **CLOSE WAIT** – состояние, соответствующее получению абонентом В запроса на завершение соединения и отправке подтверждения абоненту А. Абонент В выполняет инициацию завершения соединения и посылает запрос на разрыв соединения;

11. **LAST-ACK** – состояние, соответствующее завершению процедуры закрытия соединения (абонент В получает подтверждение запроса на завершение соединения и удаляет запись о нем).

Для формализации процесса информационного обмена [2-4] введем следующую систему определений, ограничений и допущений, которая фактически определяет границы будущего исследования. Она такова:

Определение 1. Запрос на установление соединения (syn) – служебное сообщение от одного абонента к другому абоненту с соответствующим флагом syn в заголовке.

Определение 2. Запрос на завершение соединения - служебное сообщение от одного абонента к другому абоненту с соответствующим флагом fin в заголовке.

Определение 3. Подтверждение – служебное сообщение, передаваемое после получения запроса, содержащее номер следующего ожидаемого сегмента в поле ack заголовка.

Определение 4. Номер сегмента – поле заголовка, указывающее на номер первого передаваемого октета данных (исходный номер сегмента для сообщения с флагом syn).

Определение 5. Абонент – участник информационного обмена.

Определение 6. Таймаут – промежуток времени, по истечении которого можно гарантировать, что сообщения соединения больше не перемещаются по сети.

Определение 7. Инициация соединения – процесс установления соединения, заключающийся в отправке запросов и подтверждений по правилам TCP.

Определение 8. Закрытие соединения – процесс, происходящий после информационного обмена, заключающийся в отправке запросов и подтверждений по правилам TCP.

Определение 9. Разрыв соединения – процесс немедленного прекращения информационного обмена в случае коллизии (сбой, получение неверных сегментов данных, установление нежелательного соединения).

Определение 10. Коллизия – конфликтная ситуация, характеризующаяся наличием сбоя, получения неверных сегментов данных, установления нежелательного соединения и т.п.

Определение 11. Объем передаваемой полезной нагрузки ($V_{ППН}$) – количество бит передаваемой информации.

Ограничение 1. Объем передаваемой полезной нагрузки при организации и осуществлении информационного обмена фиксирован.

Допущение 1. Каналы связи от абонента А к абоненту В и от абонента В к абоненту А нестационарные по вероятности символьной ошибки P_O.

Допущение 2. Процесс информационного обмена является «вложенным» и будет рассмотрен на следующем этапе исследования.

Физический смысл переходов временной диаграммы (представленной на рисунке 2):

Syn (x) – запрос на установление соединения, где x – номер сегмента, присваиваемый абонентом А;

Syn (y) - запрос на установление соединения, где y – номер сегмента, присваиваемый абонентом В.

Fin (a) – запрос на завершение соединения, где a – номер сегмента, присваиваемый абонентом А в зависимости от общего количества передаваемых сегментов;

Fin (b) – запрос на завершение соединения, где b – номер сегмента, присваиваемый абонентом В в зависимости от общего количества передаваемых сегментов.

Ack (x+1) – служебное сообщение, передаваемое после получения запроса syn (x), содержащее номер следующего ожидаемого сегмента (x+1) в поле ack заголовка.

Ack (y+1) – служебное сообщение, передаваемое после получения запроса syn (y), содержащее номер следующего ожидаемого сегмента (y+1) в поле ack заголовка.

Ack (a+1) – служебное сообщение, передаваемое после получения запроса fin (a), содержащее номер следующего ожидаемого сегмента (a+1) в поле ack заголовка.

Ack (b+1) – служебное сообщение, передаваемое после получения запроса fin (b), содержащее номер следующего ожидаемого сегмента (b+1) в поле ack заголовка.

Для большей визуализации исследуемого процесса представим его в виде диаграммы Хассе. Внешний вид диаграммы представлен на рисунке 3.

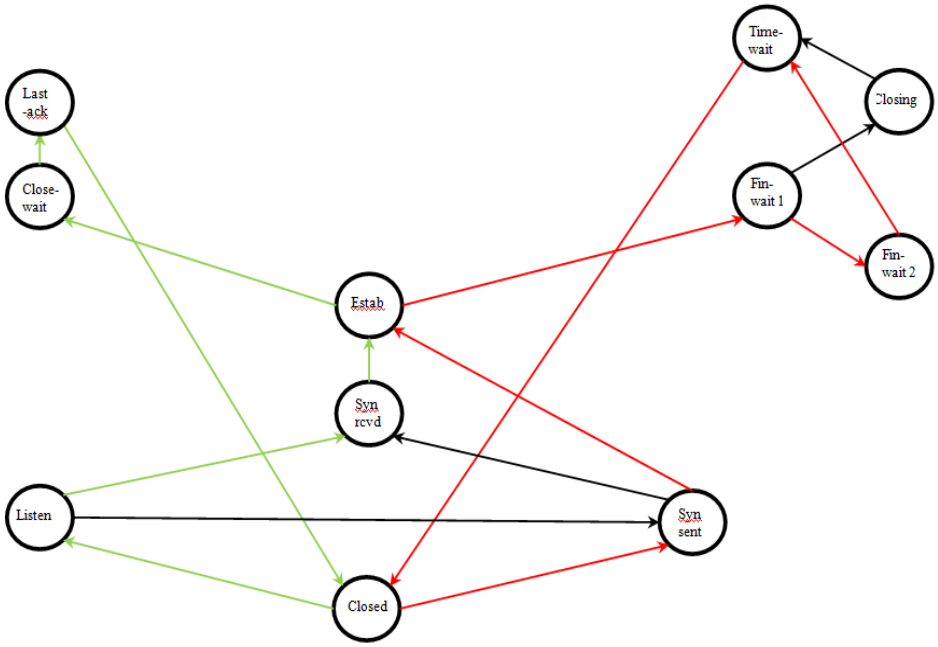


Рисунок 3 – Диаграмма Хассе процесса информационного обмена в TCP соединении

На рисунке красным цветом обозначена последовательность логических переходов для абонента А, зеленым – для абонента В. Черным цветом указаны возможные переходы в нестандартных ситуациях (принудительный разрыв соединения, попытка одновременной организации соединения абонентами и др.). Физический смысл состояний, указанных на диаграмме идентичен состояниям конечного автомата, представленного на рисунке 1.

Таким образом, представлен конечный автомат, временная диаграмма и диаграмма Хассе рассматриваемого процесса. Тем самым формализован процесс установления соединения в пакетном радиоканале не базе TCP-подобного соединения для дальнейшего построения математической модели [2-5] рассматриваемого процесса.

Литература

1. Internet Engineering Task Force (IETF) W. Eddy, Request for Comments: 9293 August 2022, ISSN: 2070-1721.

2. Порядок использования индивидуального и зондирующего вызова в системе радиосвязи второго поколения стандарта mil-std-188-141b для последующего измерения качества принятого ale сигнала [Текст] : / В. А. Цимбал, М. Ю. Попов, П. К. Калач, Р. В. Попов, Д. А. Киреев // Сборник научных работ. X Всероссийской научной конференции «Системный анализ и прикладная синергетика» (п. Нижний Архыз) – г. Ростов-на-Дону; Таганрог : Издательство Южного федерального университета, 2021. – Секция V – С. 260–268. ISBN 978-5-9275-3879-9.

3. M.Popov et al., «Synthesis of Transient Probabilities of Absorbing Final Markov Chain Describing Traffic Management Protocol Process of the Mil-std-188-141b Standard to Ensure the Probability-time Characteristic Requirements», 2020 22th International Conference on Digital Signal Processing and its Applications (DSPA), Moscow, Russia, 2020, pp. 1-4, doi: 10.1109/DSPA48919.2020.9213298.

4. Математическая модель процесса доведения сообщения по радиоканалу управления специальной АСУ в условиях энергетического и информационного воздействия на базе конечной марковской цепи с доходами [Текст] / В. А. Цимбал, М. Ю. Попов, В. В. Илющенко, Е. В. Калганов, А. М. Деркач // Доклады 24-й международной конференции «Цифровая обработка сигналов и её применение» (DSPA-2022); Серия: цифровая обработка сигналов и ее применение / Рос. науч.-техн. общ. радиотехн., электрон. и связи им. А.С. Попова. – М. : Вып. XXIV.,– С. 70–75. – ISBN 978-5-905278-49-5.

5. Программа расчета численных характеристик поглощающей конечной полумарковской цепи на базе языка программирования С++ в среде Visual studio [Текст] / М. Ю. Попов, В. Б. Девятияров, И. Б. Иванов, Д. А. Киреев, И А. Харитонов // Всероссийская конференция «Современные технологии обработки сигналов» (СТОС-2023) ; Доклады ; Серия: научные всероссийские конференции / Рос. науч.-техн. общ. радиотехн., электрон. и связи им. А.С. Попова. – М., 2023. – Вып. IV.,– С. 42–47. – ISBN 978-5-905278-55-6.

FORMALIZATION OF CONNECTION ESTABLISHMENT PROCESS AND IMPLEMENTATION OF INFORMATION EXCHANGE IN PACKET RADIO CHANNEL BASED ON TCP-LIKE CONNECTION

Dr.Sci.Tech. Popov M.Yu.[1], Cand.Tech.Sci. Kabanovich S.G.[1],
Torbenko Ya.K.[1], Popov R.V.[2]

[1]Branch of military academy of Rocket strategic forces of a name of Peter the Great
(c. Serpukhov, Moscow region).
[2]16 Central scientific research institute of the Ministry of Defence of the Russian Federation.

The article based on the state machine, time diagrams and Hasse diagram presents a formalization of the process of establishing a connection and carrying out information exchange in a packet radio channel based on a TCP-like connection. The physical meaning of the corresponding states and transitions is described. A system of necessary definitions, limitations and assumptions is given.

ФОРМАЛИЗАЦИЯ ПРОЦЕССА НАСТРОЙКИ РЕГУЛИРУЕМЫХ ПАРАМЕТРОВ ПРОТОКОЛА ПЕРЕДАЧИ ДАННЫХ ПОДУРОВНЕМ УПРАВЛЕНИЯ ТРАФИКА В ПЕРСПЕКТИВНОМ ПАКЕТНОМ ДЕКАМЕТРОВОМ РАДИОКАНАЛЕ

д.т.н., доц. Попов М.Ю.[1], к.т.н. Девятияров В.Б.[2],
асп. Киреев Д.А,[1] асп. Торбенко Я.К.[1]

[1]Филиал военной академии РВСН имени Петра Великого в г. Серпухове.
[2]АО «Корпорация «Московский институт теплотехники», г. Москва.

В статье представлены особенности процесса настройки регулируемых параметров протокола передачи данных подуровнем управления трафика в пакетном декаметровом радиоканале. Процесс формализован и представлен в виде схемы управления информационным обменом.

Особенности распространения радиоволн коротковолнового диапазона и наличие соответствующих технологий обеспечения требуемой помехоустойчивости и надежности радиосвязи в данном диапазоне длин волн открывает большие возможности по использованию его в системах передачи данных.

Одновременно с этим, существует объективная потребность в системах информационного обмена функционирующих в сложных помеховых условиях [1,2], которые, потенциально, могут быть реализованы на коротковолновых (декаметровых) радиоканалах связи.

В связи с этим научно-исследовательскими организациями и предприятиями промышленности активно ведутся работы по созданию декаметровых радиосистем и радиосетей. Особое место в данных работах уделено развитию декаметровых систем и сетей передачи данных, функционирующих в пакетном режиме с использованием технологии программно-определяемого радио.

Таким образом, работа, направленная на исследования вопросов моделирования настройки регулируемых параметров протокола передачи данных в перспективном пакетном декаметровом радиоканале общего назначения является своевременной и актуальной.

Особенности процесса настройки регулируемых параметров протокола передачи данных подуровнем управления трафика.

Известно, что все уровни ЭМ ВОС взаимосвязаны алгоритмическим, программным и другим обеспечением, по средствам соответствующих алгоритмов, процедур и правил. Одной из таких процедур является процедура настройки на сеансе связи регулируемых параметров (РП) протокола передачи данных (ППД) и управления, передачей данных по нему. Процедура обеспечивает настройку количества пакетов сообщений, переносящих информацию, объем этой информации и количество переспросов непринятых пакетов в пакетном радиоканале. Процедура реализована на уровне протокола управления трафика и выполняется подпроцессором вычислительного комплекса аппаратуры передачи данных (АПД) объекта управления и может быть оценена по средствам вероятностно-временных характеристик (ВВХ) процесса настройки.

Для нахождения таких ВВХ требуется построение соответствующей модели процесса [3-5] настройки регулируемых параметров протокола передачи данных протоколом управления трафиком.

Обобщенная схема настройки РП ППД подпроцессором вычислительного комплекса аппаратуры передачи данных (АПД) объекта управления представлена на рисунке 1. Схема поясняет процесс настройки РП ППД на сеансе связи в терминах буферов и каналов обработки (КО) вычислительного комплекса АПД.

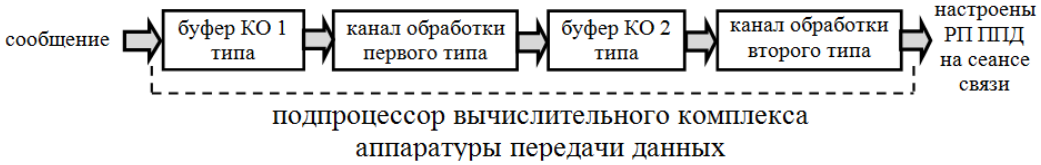


Рисунок 1 - Схема настройки РП ППД

В общем виде, процесс настройки РП ППД следующий: поступающее сообщение в начале всегда поступает в буфер КО 1-го типа, после нахождения определенного времени в нем, при отсутствии загрузки каналов обработки 1-го типа, поступает в него. Далее, после соответствующей обработки в КО 1-го типа, поступает в буфер КО 2-го типа, после нахождения в нем определенного времени и свободном КО 2-го типа, поступает в КО 2-го типа, где и осуществляется окончательная настройка РП ППД. Все этапы характеризуются соответствующими временными сроками.

Организация процесса информационного обмена в рассматриваемом декаметровом радиоканале, в части фаз процесса, в общем виде, представлена на рисунке 2.

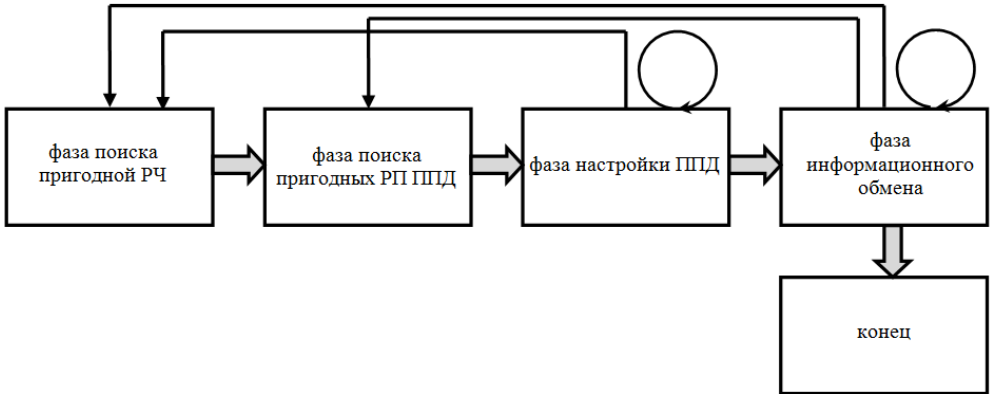


Рисунок 2 – Схема организация процесса информационного обмена в перспективном декаметровом радиоканале

Здесь и далее, к описанию поведения фаз рассматриваемого процесса будет применим следующий тезаурус: - начало; - продолжение; - прерывание; - возврат; - завершение. Для улучшения наглядности рассматриваемого процесса на рисунке 3 представлена соответствующая диаграмма Хассе.

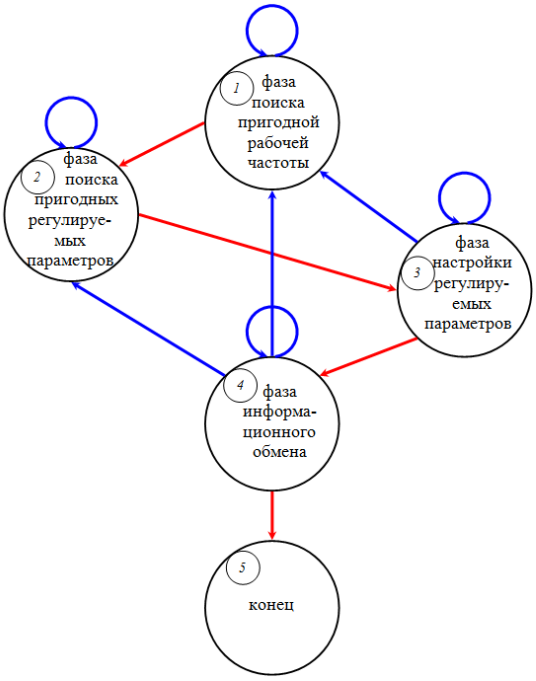


Рисунок 3 – Диаграмма Хассе процесса организации информационного обмена в перспективном декаметровом радиоканале

Красным цветом на рисунке 3 показана эволюция процесса в части начала и завершения соответствующих фаз процесса. Синим цветом значится эволюция процесса в части продолжения, прерывания и возврата к соответствующим фазам.

Рассматриваемый процесс информационного обмена, содержащий указанные фазы, функционирует в условиях, как внутреннего (от внутренних регуляторов), так и внешнего управления (от внешних, независящих, от рассматриваемого процесса, регуляторов).

На рисунке 4, на качественном уровне, представлена формализированная схема управления (внутренним и внешним процессом) информационным обменом в радиоканале.

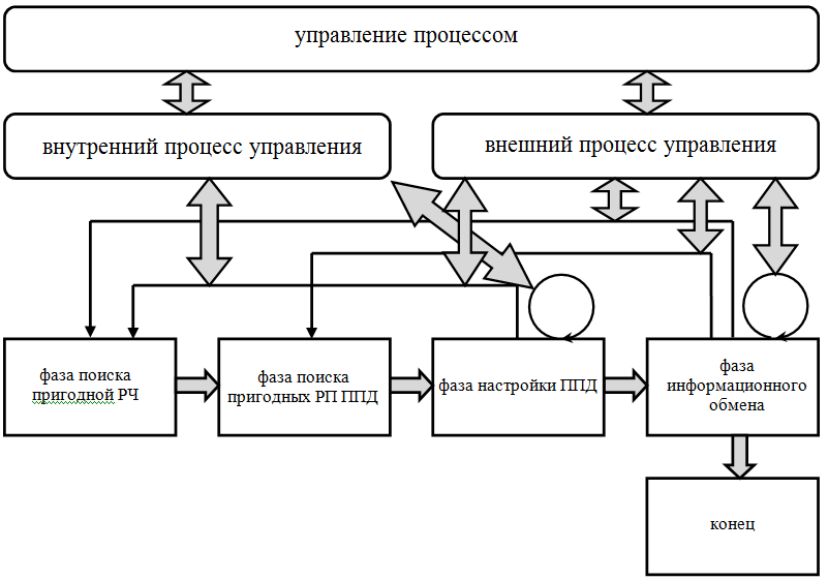


Рисунок 4 - Формализированная схема управления (внутренним и внешним процессом) информационным обменом в общем виде

Таким образом, приведены особенности процесса настройки регулируемых параметров протокола передачи данных подуровнем управления трафика в перспективном пакетном декаметровом радиоканале и соответствующая схема управления информационным обменом.

Литература

1. Регулярные конечные марковские цепи, описывающие процесс доведения сообщения объекту управления критической инфраструктуры по радиоканалу управления автоматизированной системы управления в условиях энергетического и информационного воздействия [Текст] / М. Ю. Попов, Е. В. Калганов, Н. В. Попова, Д. А. Киреев, И. Б. Иванов // Проектирование и технология электронных средств : всерос. науч.–техн. жур. – Владимир, 2023. – № 1. – С. 16–23. – ISSN 2071–9809.

2. Математическая модель процесса доведения сообщения по радиоканалу управления специальной АСУ в условиях энергетического и информационного воздействия на базе конечной марковской цепи с доходами [Текст] / В. А. Цимбал, М. Ю. Попов, В. В. Илющенко, Е. В. Калганов, А. М. Деркач // Доклады 24-й международной конференции «Цифровая обработка сигналов и её применение» (DSPA-2022); Серия: цифровая обработка сигналов и ее применение / Рос. науч.-техн. общ. радиотехн., электрон. и связи им. А.С. Попова. – М. : Вып. XXIV.,– С. 70–75. – ISBN 978-5-905278-49-5.

3. Формализация процесса обслуживания абонента системой обеспечения вызова экстренных оперативных служб по единому номеру «112» конечной полумарковской цепью [Текст] / В. А. Цимбал, М. Ю. Попов, Д. А. Киреев, Н. В. Попова, И. Б. Иванов // Проектирование и технология электронных средств : всерос. науч.–техн. жур. – Владимир, 2023. – № 2. – С. 12–20. – ISSN 2071–9809.

4. Обоснование регулируемых параметров модели процесса обслуживания абонента системой обеспечения вызова экстренных оперативных служб по единому номеру «112» [Текст] / М. Ю. Попов, В. Б. Девятияров, Д. А. Киреев, Н. В. Попова, И. Б. Иванов // Проектирование и технология электронных средств : всерос. науч.–техн. жур. – Владимир, 2023. – № 2. – С. 52–57. – ISSN 2071–9809.

5. Описание регулируемых параметров модели, созданной на базе полумарковской цепи, описывающей действия оператора системы обеспечения вызова экстренных оперативных служб по единому номеру «112» [Текст] / М. Ю. Попов, С. С. Чайков, Н. В. Попова, С. А. Кравец, И. Б. Иванов // Доклады 25-й международной конференции «Цифровая обработка сигналов и её применение» (DSPA-2023); Серия: цифровая обработка сигналов и ее применение / Рос. науч.-техн. общ. радиотехн., электрон. и связи им. А.С. Попова. – М., 2023. - Вып. XXV.,– С. 51–55. – ISBN 978-5-905278-53-2.

FORMALIZATION OF PROCESS OF ADJUSTMENT OF ADJUSTABLE PARAMETERS OF THE REPORT OF DATA TRANSMISSION BY THE SUBTOTAL OF MANAGEMENT OF THE TRAFFIC IN PERSPECTIVE PACKAGE SHORT-WAVE THE RADIO CHANNEL

Dr.Sci.Tech. Popov M.Yu.[1], Cand.Tech.Sci. Devyatiyarov V.B.[2],
Kireev D.A.[1], Torbenko Ya.K.[1]

[1]Branch of military academy of Rocket strategic forces of a name of Peter the Great
(c. Serpukhov, Moscow region).
[2]Joint-stock company corporation «Moscow institute of heat engineering».

In article features of process of adjustment of adjustable parametres of the report of data transmission by a subtotal of management of the traffic in a package short-wave radio channel are presented. Process is formalized and presented in the form of the management scheme by an information exchange.

ГРАФОВАЯ МОДЕЛЬ ПРОЦЕССА НАСТРОЙКИ РЕГУЛИРУЕМЫХ ПАРАМЕТРОВ ПРОТОКОЛА ПЕРЕДАЧИ ДАННЫХ ПОДУРОВНЕМ УПРАВЛЕНИЯ ТРАФИКА В ПЕРСПЕКТИВНОМ ПАКЕТНОМ ДЕКАМЕТРОВОМ РАДИОКАНАЛЕ

д.т.н., доц. Попов М.Ю.[1], к.т.н. Шаблюк С.М.[2],
асп. Торбенко Я.К.[1], Попова Н.В.[1]

[1]Филиал военной академии РВСН имени Петра Великого в г. Серпухове.
[2]АО НПО «Импульс», г. Санкт-Петербург.

В статье представлен граф состояний и переходов конечной полумарковской цепи описывающий процесс настройки регулируемых параметров протокола передачи данных подуровнем управления трафика в перспективном пакетном декаметровом радиоканале общего назначения функционирующего в условиях наличия перерывов радиосвязи. Описан физический смысл его состояний, переходных вероятностей и шагов перехода.

В силу особенностей рассматриваемого процесса настройки регулируемых параметров протокола передачи данных подуровнем управления трафика модель, возможно, построить на базе научно-методического аппарата конечных цепей Маркова [1-3].

Соответствующие графы состояний и переходов (ГСП) с размеченными переходными вероятностями (рисунок 1 а)) и шагами переходов (рисунок 1 б)) представлены на рисунке 1.

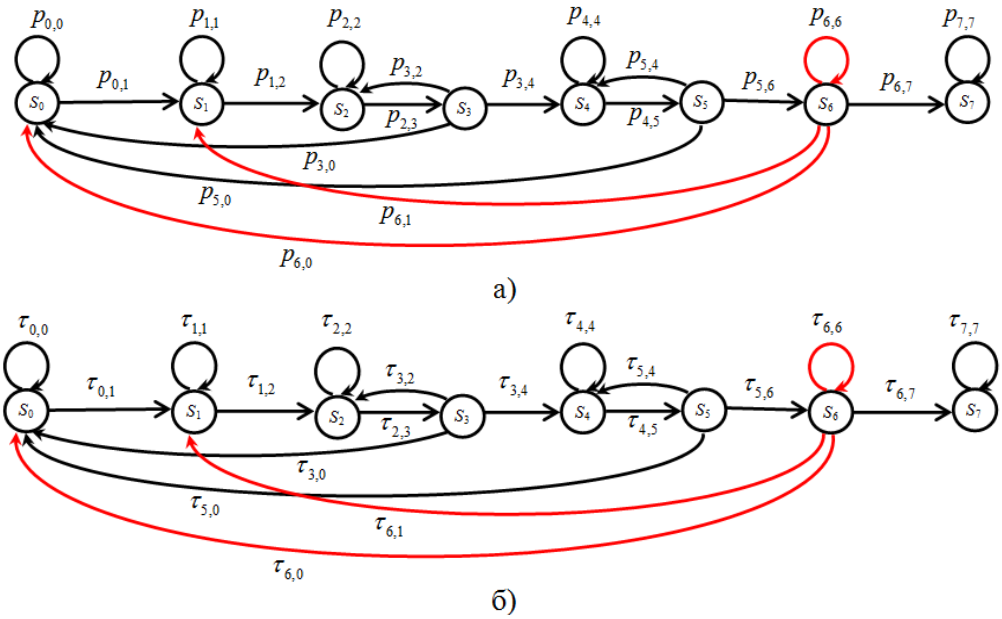


Рисунок 1 – ГСП КПЦ марковской цепи процесса настройки РП ППД в пакетном декаметровом радиоканале подуровнем управления трафиком в условиях наличия перерывов радиосвязи

Физический смысл состояний ГСП цепи описывающий рассматриваемый процесс, с учетом его фаз, таков:

S_0 – состояние, соответствующие осуществлению фазы поиска пригодной рабочей частоты (РЧ) – осуществляется анализ качества радиоканала связи с целю нахождения пригодной РЧ;

S_1 – состояние, соответствующие осуществлению фазы поиска пригодных регулируемых параметров (РП) протокола передачи данных (ППД) - на РЧ, определенной в ходе фазы поиска пригодной РЧ, определяются текущие регулируемые параметры ППД, формируется служебное сообщение содержащие найденные их значения;

S_2 – состояние, соответствующие началу фазы настройки РП ППД - поступление сообщения в буфер канала обработки (КО) 1 типа;

S_3 – состояние, соответствующие продолжению фазы настройки РП ППД - поступление сообщения в КО 1 типа;

S_4 – состояние, соответствующие продолжению фазы настройки РП ППД - поступление сообщения в буфер КО 2 типа;

S_5 – состояние, соответствующие продолжению фазы настройки РП ППД - поступление сообщения в КО 2 типа;

S_6 – состояние, соответствующие осуществлению фазы информационного обмена - работает ППД с текущими РП;

S_7 – состояние, соответствующие окончанию фазы информационного обмена - вся передаваемая информация доставлена, конец.

На рисунке 2 представлен ГСП с размеченными фазами процесса.

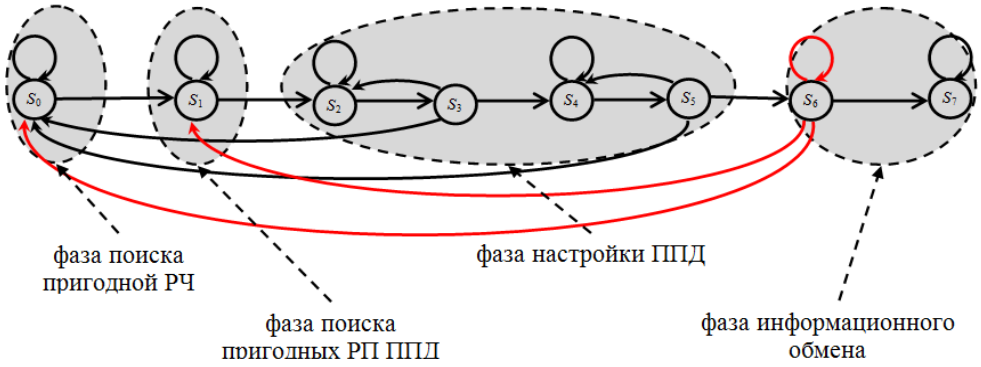


Рисунок 2 - ГСП с размеченными фазами процесса

В силу физических особенностей процесса, переходы из состояния в состояние представленного графа будут характеризоваться не только вероятностями (переходными вероятностями), но и соответствующими временами (шагами переходов) [4,5].

Переходные вероятности и шаги переходов ГСП имеют следующий физический смысл:

$p_{0,0}$; $\tau_{0,0}$ - вероятность и время продолжения фазы поиска пригодной РЧ - анализ качества радиоканала связи с целью нахождения пригодной РЧ;

$p_{0,1}$; $\tau_{0,1}$ - вероятность и время завершения фазы поиска пригодной РЧ - вероятность и время определения пригодной РЧ – вероятность и время начала фазы поиска пригодных РП ППД;

$p_{1,1}$; $\tau_{1,1}$ - вероятность и время продолжения фазы поиска пригодных РП ППД – вероятность и время поиска пригодных регулируемых параметров;

$p_{1,2}$; $\tau_{1,2}$ - вероятность и время завершения фазы поиска пригодных РП ППД - вероятность и время определения пригодных регулируемых параметров ППД, формирования и отправка служебного сообщения в буфер обработки 1 типа содержащего их значения - вероятность и время начала фазы настройки ППД;

$p_{2,2}$; $\tau_{2,2}$ - вероятность и время продолжения фазы настройки ППД - вероятность и время нахождения служебного сообщения в буфере канала обработки (КО) 1 типа;

$p_{2,3}$; $\tau_{2,3}$ - вероятность и время продолжения фазы настройки ППД - вероятность и время поступления служебного сообщения в КО 1 типа;

$p_{3,2}$; $\tau_{3,2}$ - вероятность и время продолжения фазы настройки ППД - вероятность и время возникновения признака задержки при обработки сообщения в КО 1 типа;

$p_{3,0}$; $\tau_{3,0}$ - вероятность и время прерывания фазы настройки ППД - вероятность и время возникновения признака конфликта РП - вероятность и время возврата к фазе поиска пригодной РЧ;

$p_{3,4}$; $\tau_{3,4}$ - вероятность и время продолжения фазы настройки ППД - вероятность и время поступления служебного сообщения в буфер канала обработки 2 типа;

$p_{4,4}$; $\tau_{4,4}$ - вероятность и время продолжения фазы настройки ППД - вероятность и время нахождения служебного сообщения в буфере КО 2 типа;

$p_{4,5}$; $\tau_{4,5}$ - вероятность и время продолжения фазы настройки ППД - вероятность и время поступления служебного сообщения в КО 2 типа;

$p_{5,0}$; $\tau_{5,0}$ - вероятность и время прерывания фазы настройки ППД - вероятность и время возникновения признака внешнего прерывания процесса - вероятность и время возврата к фазе поиска пригодной РЧ;

$p_{5,4}$; $\tau_{5,4}$ - вероятность и время продолжения фазы настройки ППД - вероятность и время возникновения признака задержки при обработке сообщения в КО 2 типа;

$p_{5,6}$; $\tau_{5,6}$ - вероятность и время завершения фазы настройки ППД - вероятность и время обработки сообщения в КО 2 типа завершена - вероятность и время начала фазы информационного обмена;

$p_{6,6}$; $\tau_{6,6}$ - вероятность и время продолжения фазы информационного обмена - вероятность и время появления перерыва радиосвязи 1 типа;

$p_{6,1}$; $\tau_{6,1}$ - вероятность и время прерывания фазы информационного обмена - вероятность и время появления перерыва радиосвязи 2 типа - вероятность и время возврата к фазе поиска пригодных РП ППД;

$p_{6,0}$; $\tau_{6,0}$ - вероятность и время прерывания фазы информационного обмена - вероятность и время появления перерыва радиосвязи 3 типа - вероятность и время возврата к фазе поиска пригодной РЧ;

$p_{6,7}$; $\tau_{6,7}$ - вероятность и время завершения фазы информационного обмена - вероятность и время доведения всей передаваемой информации;

$p_{7,7}$; $\tau_{7,7}$ - вероятность и время перехода поглощающего состояния цепи.

Сформируем матрицу переходных вероятностей (МПВ) и матрицу шагов перехода (МШП) представленной поглощающей цепи. МПВ имеет вид:

$$P_{[8,8]} = \begin{Vmatrix} p_{0,0} & p_{0,1} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & p_{1,1} & p_{1,2} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & p_{2,2} & p_{2,3} & 0 & 0 & 0 & 0 \\ p_{3,0} & 0 & p_{3,2} & 0 & p_{3,4} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & p_{4,4} & p_{4,5} & 0 & 0 \\ p_{5,0} & 0 & 0 & 0 & p_{5,4} & 0 & p_{5,6} & 0 \\ p_{6,0} & p_{6,1} & 0 & 0 & 0 & 0 & p_{6,6} & p_{6,7} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & p_{7,7} \end{Vmatrix}. \quad (1)$$

МШП имеет вид:

$$T_{[8,8]} = \begin{Vmatrix} \tau_{0,0} & \tau_{0,1} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \tau_{1,1} & \tau_{1,2} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \tau_{2,2} & \tau_{2,3} & 0 & 0 & 0 & 0 \\ \tau_{3,0} & 0 & \tau_{3,2} & 0 & \tau_{3,4} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \tau_{4,4} & \tau_{4,5} & 0 & 0 \\ \tau_{5,0} & 0 & 0 & 0 & \tau_{5,4} & 0 & \tau_{5,6} & 0 \\ \tau_{6,0} & \tau_{6,1} & 0 & 0 & 0 & 0 & \tau_{6,6} & \tau_{6,7} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \tau_{7,7} \end{Vmatrix}. \quad (2)$$

Таким образом, на базе поглощающей конечной полумарковской цепи создана математическая модель процесса настройки регулируемых параметров протокола передачи данных подуровнем управления трафика в перспективном пакетном декаметровом радиоканале

в условиях наличия перерывов радиосвязи подпроцессором вычислительного комплекса аппаратуры передачи данных. Описан физический смысл состояний соответствующего графа состояний и переходов цепи, его переходных вероятностей и шагов перехода.

Литература

1. Описание регулируемых параметров модели, созданной на базе полумарковской цепи, описывающей действия оператора системы обеспечения вызова экстренных оперативных служб по единому номеру «112» [Текст] / М. Ю. Попов, С. С. Чайков, Н. В. Попова, С. А. Кравец, И. Б. Иванов // Доклады 25-й международной конференции «Цифровая обработка сигналов и её применение» (DSPA-2023); Серия: цифровая обработка сигналов и ее применение / Рос. науч.-техн. общ. радиотехн., электрон. и связи им. А.С. Попова. – М., 2023. - Вып. XXV.,– С. 51–55. – ISBN 978-5-905278-53-2.

2. Регулярные конечные марковские цепи, описывающие процесс доведения сообщения объекту управления критической инфраструктуры по радиоканалу управления автоматизированной системы управления в условиях энергетического и информационного воздействия [Текст] / М. Ю. Попов, Е. В. Калганов, Н. В. Попова, Д. А. Киреев, И. Б. Иванов // Проектирование и технология электронных средств : всерос. науч.–техн. жур. – Владимир, 2023. – № 1. – С. 16–23. – ISSN 2071–9809.

3. Математическая модель процесса доведения сообщения по радиоканалу управления специальной АСУ в условиях энергетического и информационного воздействия на базе конечной марковской цепи с доходами [Текст] / В. А. Цимбал, М. Ю. Попов, В. В. Илющенко, Е. В. Калганов, А. М. Деркач // Доклады 24-й международной конференции «Цифровая обработка сигналов и её применение» (DSPA-2022); Серия: цифровая обработка сигналов и ее применение / Рос. науч.-техн. общ. радиотехн., электрон. и связи им. А.С. Попова. – М. : Вып. XXIV.,– С. 70–75. – ISBN 978-5-905278-49-5.

4. Формализация процесса обслуживания абонента системой обеспечения вызова экстренных оперативных служб по единому номеру «112» конечной полумарковской цепью [Текст] / В. А. Цимбал, М. Ю. Попов, Д. А. Киреев, Н. В. Попова, И. Б. Иванов // Проектирование и технология электронных средств : всерос. науч.–техн. жур. – Владимир, 2023. – № 2. – С. 12–20. – ISSN 2071–9809.

5. Обоснование регулируемых параметров модели процесса обслуживания абонента системой обеспечения вызова экстренных оперативных служб по единому номеру «112» [Текст] / М. Ю. Попов, В. Б. Девятияров, Д. А. Киреев, Н. В. Попова, И. Б. Иванов // Проектирование и технология электронных средств : всерос. науч.–техн. жур. – Владимир, 2023. – № 2. – С. 52–57. – ISSN 2071–9809.

THE COUNT OF MODEL OF PROCESS OF ADJUSTMENT OF ADJUSTABLE PARAMETERS OF THE REPORT OF DATA TRANSMISSION BY THE SUBTOTAL OF MANAGEMENT OF THE TRAFFIC IN PERSPECTIVE PACKAGE SHORT-WAVE THE RADIO CHANNEL

Dr.Sci.Tech. Popov M.Yu.[1], Cand.Tech.Sci. Shablyuk S.M.[2],
Torbenko Ya.K.[1], Popova N.V.[1]

[1]Branch of military academy of Rocket strategic forces of a name of Peter the Great
(c. Serpukhov, Moscow region).
[2]Research-and-production association «Impulse», St.-Petersburg.

In article the state graph and transitions final semimarkovsky is presented a chain describing process of adjustment of adjustable parametres of the report of data transmission by a subtotal of management of the traffic in a perspective package short-wave radio channel of a general purpose functioning in the conditions of presence of breaks of a radio communication. The physical sense of its conditions, transitive probabilities and transition steps is described.

ОПТИМИЗАЦИЯ ИНТЕРПОЛЯТОРА НА ОСНОВЕ ФИЛЬТРА ФАРРОУ

асп. Буссе А.А.[1], преп. к.т.н. Соловьев Д.М.[1], студ. Мещеряков А.А.[1]

[1] Московский физико-технический институт (НИУ)

В данной статье рассматривается оптимизация структуры и коэффициентов фильтров Фарроу применительно к задаче интерполяции. На данный момент существующие подходы либо рассматривают фильтры Фарроу как систему с перестраиваемой задержкой VFD (Variable Fractional Delay) либо как систему, восстанавливающую непрерывный сигнал. На практике, зачастую задача ставится как интерполяция входного сигнала в рациональное число раз и количество известных заранее конфигураций ограничено. В работе рассматривается данная задача и предлагаются методы для поиска оптимального решения. Так же исследования фильтров Фарроу показали, что замена полиномиального базиса позволяет улучшить решение, поэтому предлагаемые решения представлены в общем виде и могут быть использованы с любым базисом, в частности для сравнения был реализован базис из комплексных экспонент.

Введение

Задача интерполяции дискретного сигнала является распространенной в области цифровой обработки сигналов и возникает как в задачах аудио обработки, так и в задачах формирования и приема радиосигналов. Для данной задачи необходим фильтр с дробной задержкой. Данный фильтр характеризуется тем, что имеет практически постоянную групповую задержку τ в полосе сигнала. Во многих реальных сценариях значение задержки должно меняться. В таком случае получается, что значение коэффициентов фильтра $\mathbf{h}$ так же должны быть функциями от задержки $\mathbf{h}(\tau)$. Наиболее удобной и эффективной структурой, реализующей подобный фильтр, значения коэффициентов которого являются функцией переменной, является фильтр Фарроу [1]. Структура данного фильтра представлена на рисунке (1а) и включает набор из M подфильтров $\mathbf{c}_m, m \in [0, \dots, M-1]$ порядка $N-1$ и схему Горнера для вычисления полинома от τ. Набор всех коэффициентов будем представлять в виде матрицы $N \times M$, где коэффициенты подфильтров являются столбцами $\mathbf{C} = \left(c_{n,m}\right)_{0 \le n \le N-1, 0 \le m \le M-1} = [\mathbf{c}_0, \mathbf{c}_1, \dots, \mathbf{c}_{M-1}]$.

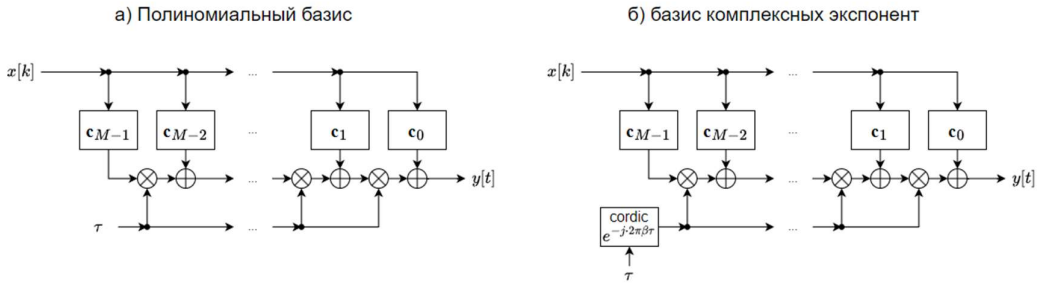


Рисунок 1 – Структура фильтра Фарроу

Зависимость коэффициентов фильтра от τ можно выразить следующим образом $\mathbf{h}(\tau) = \sum_{m=0}^{M-1} \mathbf{c}_m \tau^m$. Данную зависимость можно обобщить $\mathbf{h}(\tau) = \left(h_n(\tau)\right)_{0 \le n \le N-1} = \sum_{m=0}^{M-1} \mathbf{c}_m a_m(\tau) = \mathbf{C}\mathbf{a}(\tau)$, где $\mathbf{a}(\tau) = \left(a_m(\tau)\right)_{0 \le m \le M-1}$ – базисные функции. Оригинальная структура Фарроу использует полиномиальный базис $a_m(\tau) = \tau^m$. Данный базис имеет преимущество в том, что может быть эффективно реализован через схему Горнера. Так же для задач дробной задержки были исследованы другие базисы [2,3,4]. Наиболее интересен базис, составленный из комплексных экспонент [2], позволяющий значительно улучшить характеристика фильтра (например подавление), тем самым потенциально сократить минимальное необходимое количество подфильтров M, но взамен придется заменить схему Горнера на более сложную, изображенную на рисунке (1б). Данный базис далее будем называть гармоническим. Стоит отметить, что фильтр с базисом из комплексных экспонент, хоть и описывается комплексными переменными, но вследствие симметрий его подфильтры реализуются тем же количеством

действительных умножений, а замена схемы Горнера требует примерно вдвое больше действительных умножений.

Коэффициенты матрицы $\mathbf{C}$ являются объектом для оптимизации. Критерий оптимизации зависит от постановки задачи. Для фильтров Фарроу наиболее исследованной является задача оптимизации VFD (Variable Fractional Delay) [1-7]. В данной задаче необходимо приблизить фильтр к идеальному фильтру дробной задержки в полосе сигнала для всех требуемых задержек. В данной статье мы исследуем задачу передискретизации, в которой требуется изменить частоту дискретизации входного сигнала. В статье рассмотрен только случай интерполяции, но данную постановку легко расширить на случаи произвольной передискретизации путем децимации выходного сигнала.

Задача передискретизации с использованием фильтра Фарроу была исследована в работах [8,9,10,11]. Данными авторами рассматривается задача восстановления непрерывного сигнала $y(t)$ по его дискретным отсчетам $x(k)$. В данной работе мы рассматриваем так же задачу передискретизации, но в отличие от предыдущих работ нас интересует интерполяция в заданное количество раз L, которое во время работы может принимать различные значения из множества $\{L_0, \dots, L_{K-1}\}$. Такая задача распространена на практике вследствие того, что зачастую все возможные значения L известны заранее.

Интерполяция

Рассмотрим интерполяцию в заданное количество раз L. Покажем, что в таком случае фильтр Фарроу будет эквивалентен интерполяции с помощью КИХ (конечной импульсной характеристикой) фильтра порядка $LN-1$. Пусть на вход системы подается сигнал $x(k)$, а сигнал на выходе системы обозначим $y(t)$. Рассмотрим интерполяцию с помощью фильтра с коэффициентами $\mathbf{h}^{(i)} = \left(h_n^{(i)}\right)_{0 \le n \le LN-1}$ порядка $LN-1$ (индекс (i) сокращение от interpolator добавлен для отличия от импульсной характеристики фильтра Фарроу $\mathbf{h}(\tau)$). Сначала дополним сигнал $x(k)$ нулями и получим сигнал $v(k)$. Затем фильтруем сигнал $v(k)$, чтобы избавиться от изображений сигнала вне полосы. Сигнал $y(t)$ можно подразделить на L подпоследовательностей (1).

$$v(k) = \begin{cases} x\left(\frac{k}{L}\right) \text{ если } mod(k, L) = 0 \\ 0 \text{ в остальных случаях} \end{cases}$$

$$y(t) = \sum_{n=0}^{NL-1} v(n) h_{t-n}^{(i)} = \sum_{n=0}^{N-1} x(n) h_{t-nL}^{(i)}$$

$$y(Lk+l) = \sum_{n=0}^{N-1} x(n) h_{L(k-n)+l}^{(i)}, l \in [0, \dots, L-1] \quad (1)$$

Рассмотрим работу фильтра Фарроу. Выходной сигнал $y(t)$ можно неформально записать следующим образом $y(t) = x\left(\frac{t-t_0}{L}\right)$ или $y(Lk+l) = x\left(k + \frac{l-t_0}{L}\right), l \in [0, \dots, L-1]$ – как сигнал $x(n)$ с L различными задержками. Константа t_0 может быть выбрана произвольно, но логичнее всего выбрать $t_0 = (L-1)/2$, что в дальнейшем приведет к симметрии. Таким образом формально можно записать следующее. Сформируем L подпоследовательностей $y[Lk+l], l \in [0, \dots, L-1]$ путем задерживания сигнала $x(k)$ на задержки $\tau_l = \frac{l-(L-1)/2}{L}$.

$$y(Lk+l) = \sum_{n=0}^{N-1} x(n) h_{k-n}(\tau_l), l \in [0, \dots, L-1] \quad (2)$$

Из выражений (1) и (2) можно сделать вывод что работающий предложенным образом фильтр Фарроу эквивалентен интерполирующему фильтру с импульсной характеристикой $h_{Ln+l}^{(i)} = h_n(\tau_l) = \sum_{m=0}^{M-1} c_{n,m} a_m(\tau_l)$, $l \in [0, \dots, L-1]$, $n \in [0, \dots, N-1]$. Частотная характеристика $H(\omega)$ полученного фильтра записывается следующим образом:

$$H(\omega) = e^{-jwD} \sum_{l=0}^{L-1} \boldsymbol{\omega}_l^T \mathbf{C} \mathbf{a}_l$$

$$\boldsymbol{\omega}_l(\omega) = \left[e^{-j\omega(l-D)}, e^{-j\omega(L+l-D)}, \dots, e^{-j\omega(L(N-1)+l-D)}\right]^T$$

$$\mathbf{a}_l = \mathbf{a}(\tau_l) - \text{значения базисных функций в точке } \tau_l$$

$$D = \frac{LN-1}{2} - \text{задерка фильтра порядка } LN-1$$

В большинстве случаев желательно, чтобы фильтр обладал линейной ФЧХ (фазо-частотная характеристика). Для этого импульсная характеристика $\mathbf{h}^{(i)}$ должна быть симметричной. Поэтому мы опустим множитель $e^{-j\omega D}$ и оставим только действительную часть несущую информацию об АЧХ (амплитудно-частотная характеристика).

$$H(\omega) = \sum_{l=0}^{L-1} \boldsymbol{\omega}_l^T \mathbf{C}\mathbf{a}_l$$

Рассмотрим две самые распространенные задачи оптимизации: минимизация интеграла квадрата отклонения J_2 и минимизация максимума отклонения J_∞. В постановке задачи так же должны быть заданы $H_d(\omega)$ – желаемая идеальная АЧХ и $W(\omega)$ – взвешивающая функция.

$$J_2 = \int_0^{2\pi} W(\omega)|H(\omega) - H_d(\omega)|^2 d\omega$$

$$J_\infty = \max_{\omega\in[0,2\pi)} \left|W(\omega)\big(H(\omega) - H_d(\omega)\big)\right|$$

Рассмотрим задачу минимизации J_2. Для этого предположим, что $H_d(\omega)$ и $W(\omega)$ являются постоянными H_{d_r} и W_r на интервалах $[\omega_r, \omega_{r+1}), r \in [0, \dots, R-1]$.

$$J_2 = \int_0^{2\pi} W(\omega)H(\omega)^2 d\omega - 2\int_0^{2\pi} W(\omega)H(\omega)H_d(\omega)d\omega + \int_0^{2\pi} W(\omega)H_d(\omega)^2 d\omega = J_{2,0} + J_{2,1} + J_{2,2}$$

Раскроем $H(\omega)$ в $J_{2,0}$, учтем что $H(\omega), \mathbf{a}_l, \mathbf{C}$ – действительные функции/коэффициенты и соответственно $H(\omega) = \sum_{l=0}^{L-1} \boldsymbol{\omega}_l^T \mathbf{C}\boldsymbol{a}_l = \sum_{l=0}^{L-1} \boldsymbol{\omega}_l^H \mathbf{C}\mathbf{a}_l = \sum_{l=0}^{L-1} \mathbf{a}_l^T \mathbf{C}^T \boldsymbol{\omega}_l$. А также учтем, что результат $J_{2,0}$ тоже действительный и соответственно под интегралом можно оставить только действительную часть.

$$J_{2,0} = \int_0^{2\pi} W(\omega)H(\omega)^2 d\omega = \int_0^{2\pi} W(\omega) \sum_{l_1,l_2=0}^{L-1} \left(\mathbf{a}_{l_1}^T \mathbf{C}^T \boldsymbol{\omega}_{l_1}\right)\left(\boldsymbol{\omega}_{l_2}^H \mathbf{C}\mathbf{a}_{l_2}\right) d\omega = \sum_{l_1,l_2=0}^{L-1} \mathbf{a}_{l_1}^T \mathbf{C}^T \boldsymbol{\Omega}_{l_1,l_2} \mathbf{C}\mathbf{a}_{l_2}$$

$$\boldsymbol{\Omega}_{l_1,l_2} = \left(\Omega_{l_1,l_2 n,m}\right)_{0\le n,m\le N-1} = \int_0^{2\pi} W(\omega) Re\left(\boldsymbol{\omega}_{l_1}\boldsymbol{\omega}_{l_2}^H\right) d\omega$$

$$\Omega_{l_1,l_2 n,m} = \sum_{r=0}^{R-1} W_r \left(\omega_{r+1} sinc\left(\frac{\omega_{r+1}(L(n-m)+l_1-l_2)}{\pi}\right) - \omega_r sinc\left(\frac{\omega_r(L(n-m)+l_1-l_2)}{\pi}\right)\right)$$

$$sinc(x) = \frac{\sin(\pi x)}{\pi x}$$

Аналогично поступим с $J_{2,1}$ и $J_{2,2}$:

$$J_{2,1} = -2\int_0^{2\pi} W(\omega)H(\omega)H_d(\omega)d\omega = -2\sum_{l=0}^{L-1} \mathbf{a}_l^T \mathbf{C}^T \int_0^{2\pi} W(\omega)H_d(\omega)Re(\boldsymbol{\omega}_l)d\omega = -2\sum_{l=0}^{L-1} \mathbf{a}_l^T \mathbf{C}^T \mathbf{d}_l$$

$$\mathbf{d}_l = \left(d_{l_n}\right)_{0\le n\le N-1} = \int_0^{2\pi} W(\omega)H_d(\omega)Re(\boldsymbol{\omega}_l)d\omega$$

$$d_{l_n} = \sum_{r=0}^{R-1} W_r H_{d_r} \left(\omega_{r+1} sinc\left(\frac{\omega_{r+1}(Ln+l-D)}{\pi}\right) - \omega_r sinc\left(\frac{\omega_r(Ln+l-D)}{\pi}\right)\right)$$

$$J_{2,2} = \sum_{r=0}^{R-1} W_r H_{d_r}(\omega_{r+1} - \omega_r)$$

Учитывая, что $\boldsymbol{\Omega}_{k,l} = \boldsymbol{\Omega}_{l,k}^T$, запишем условие минимума ошибки $\frac{\partial J_2}{\partial \mathbf{C}} = 0$.

$$\frac{\partial J_{2,0}}{\partial \mathbf{C}} = 2\sum_{l_1,l_2=0}^{L-1} \boldsymbol{\Omega}_{l_1,l_2} \mathbf{C}\mathbf{a}_{l_1} \mathbf{a}_{l_2}^T$$

$$\frac{\partial J_{2,1}}{\partial \mathbf{C}} = -2\sum_{l=0}^{L-1} \mathbf{d}_l \mathbf{a}_l^T$$

$$\frac{\partial J_2}{\partial \mathbf{C}} = 2\sum_{l_1,l_2=0}^{L-1} \mathbf{\Omega}_{l_1,l_2}\mathbf{C}\mathbf{a}_{l_1}\mathbf{a}_{l_2}^T - 2\sum_{l=0}^{L-1}\mathbf{d}_l\mathbf{a}_l^T = 0 \quad (3)$$

Решение полученной СЛАУ (системы линейных алгебраических уравнений) (3) относительно искомых переменных $\mathbf{C}$ даст оптимальный по критерию наименьших квадратов набор коэффициентов.

Для критерия минимизации J_∞ задачу можно сформулировать как задачу линейного программирования аналогичную задаче по оптимизации фильтра [12]:

$$\text{Переменные}$$
$$c_{n,m} - \text{искомые коэффициенты}$$
$$\text{Ограничения (задаются на достаточно плотной сетке частот)}$$
$$\big(H(\omega) - H_d(\omega)\big)W(\omega) \le \delta$$
$$\big(H(\omega) - H_d(\omega)\big)W(\omega) \ge -\delta$$
$$minimize\ \delta$$

Реконфигурируемый интерполятор

Задачу перестраиваемого интерполятора можно задать следующим образом. Заданы K конфигураций. Для каждой конфигурации заданы L_k – порядок интерполяции, $H_{d_k}(\omega)$ – требуемая АЧХ, $W_k(\omega)$ – взвешивающая функция. Необходимо найти коэффициенты фильтра $\mathbf{C}$, такие что ошибка J_2 или J_∞ (в зависимости от критерия) была минимальной.

$$J_2 = \sum_{k=0}^{K-1}\int_0^{2\pi} W_k(\omega)\left|H_k(\omega) - H_{d_k}(\omega)\right|^2 d\omega$$
$$J_\infty = \max_{w\in[0,2\pi),k\in[0,\dots,K-1]}\left|W_k(\omega)\left(H_k(\omega) - H_{d_k}(\omega)\right)\right|$$
$$H_k(\omega) = \sum_{l=0}^{L_k-1}\boldsymbol{\omega}_l^T\mathbf{C}\mathbf{a}_{l,k}$$
$$\mathbf{a}_{l,k} = \mathbf{a}(\tau_{l,k})$$
$$\tau_{l,k} = \frac{l-(L_k-1)/2}{L_k}$$

Можно заметить, что $J_2 = \sum_{k=0}^{K-1} J_{2_k}$, где J_{2_k} – квадрат ошибки для k-той конфигурации. Поэтому $\frac{\partial J_2}{\partial \mathbf{C}} = \sum_{k=0}^{K-1}\frac{\partial J_{2_k}}{\partial \mathbf{C}} = 0$ так же представляет аналогичную СЛАУ относительно искомых коэффициентов $\mathbf{C}$.

В задаче минимизации J_∞ теперь ограничения задаются для всех конфигураций, а не только для одной.

$$\left(H_k(\omega) - H_{d_k}(\omega)\right)W_k(\omega) \le \delta$$
$$\left(H_k(\omega) - H_{d_k}(\omega)\right)W_k(\omega) \ge -\delta$$
$$\forall k \in [0,\dots,K-1]$$
$$minimize\ \delta$$

Результаты

Для сравнения различных алгоритмов оптимизации будем оптимизировать фильтр с $K = 3$ конфигурациями $L_0, L_1, L_2 = 2,3,5$. Требуемая АЧХ и весовая функция будут задаваться следующим образом, что бы фильтр имел необходимое усиление и подавление было одинаковым по отношению к усилению в полосе пропускания для всех конфигураций.

$$H_{d_k}(f) = \begin{cases} L_k \text{ при } -B/2L_k \le f < B/2L_k \text{ (в полосе сигнала)} \\ 0 \text{ иначе} \end{cases}$$

$$W_k(f) = \begin{cases} 1/L_k \text{ при } l/L_k - B/2L_k \le f < l/L_k + B/2L_k,\ l \in [0,\dots,\lfloor L_k/2\rfloor] \\ 0 \text{ иначе} \end{cases}$$

$$B = 0.5 - \text{полоса сигнала в нормированной частоте}$$

Для сравнения были реализованы следующие алгоритмы: алгоритмы, предложенные в данной статье с использованием полиномиального базиса (CONF_POLY) и базиса комплексных

экспонент (CONF_EXP) и алгоритмы из работы [9] (CONT_POLY), а также их модификации с использованием базиса комплексных экспонент (CONT_EXP). Оптимизация с использованием базиса из комплексных экспонент производилась по аналогии с тем, как это было предложено в работе [2].

АЧХ фильтров, полученных разными методами для параметров $N = 10, M = 4$ приведены на рисунке (Рисунок *2*), а оптимальные коэффициенты приведены в приложении.

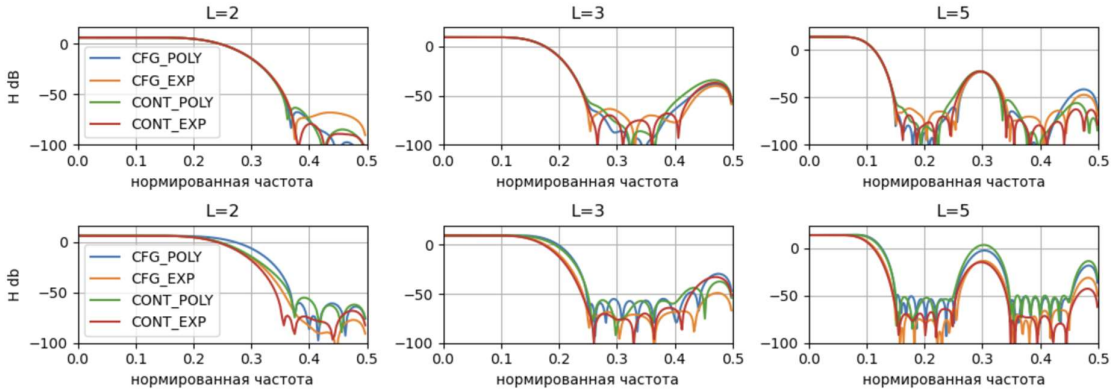


Рисунок 2: АЧХ полученных фильтров для параметров $N = 10$, $M = 4$ в разных конфигурациях $L = 2, 3, 5$. Верхние графики – фильтры, оптимизированные для минимизации J_2, нижние - J_∞.

Более детально рассмотрим зависимость J_2 и J_∞ от M для разных алгоритмов. Для этого всеми предложенными способами были оптимизированы фильтры с $N = 10$ и $M \in [2, \dots, 7]$. Результаты представлены на рисунке (3). Из полученных зависимостей видно, что $\log(J)$ убывает практически линейно пока не достигает насыщения. Насыщение обосновано максимальным подавлением, которое можно получить фильтром порядка $L_k N$, которому эквивалентен фильтр Фарроу. При использовании экспоненциального базиса ошибка убывает быстрее чем при использовании полиномиального. Можно отметить, что предложенные методы позволяют получить лучшее решение при минимизации J_2. В задаче же минимизации J_∞ решения для полиномиального базиса практически совпали, а для экспоненциального базиса разница по $20\log_{10}(J_\infty)$ была менее 1. Таким образом, приведенный метод позволяет либо улучшить решения для приведенных критериев, что наблюдается для J_2, либо показать, что текущее решение является близким к оптимальному. Это важно, т.к. благодаря улучшению характеристик становится возможным сократить необходимые для реализации фильтра ресурсы. Например предположим, что нам необходимо создать фильтр с $10 * \log_{10}(J_2) \leq -75$. В таком случае для создания фильтра с использованием экспоненциального базиса и предложенного метода оптимизации минимальным необходимым значением M будет $M_{min} = 4$, во всех других рассмотренных случаях $M_{min} = 5$.

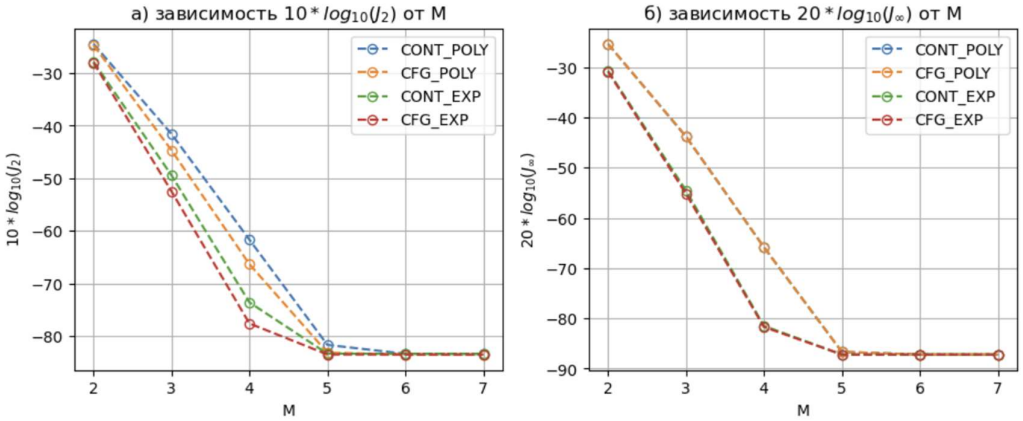


Рисунок 3. Зависимость J_2(а) и J_∞(б) от M (в логарифмическом масштабе).

Заключение

Была рассмотрена оптимизация фильтров Фарроу для нового критерия – оптимального реконфигурируемого интерполятора на основе фильтра Фарроу. Для данной задачи предложены оптимальные методы поиска коэффициентов для критериев минимизации интеграла квадрата ошибки J_2 и максимального отклонения J_∞. Предложенные методы могут быть использованы не только с полиномиальными базисными функциями, но и с любыми другими. Для сравнения были реализованы предложенные методы для полиномиального базиса и базиса комплексных экспонент. Было показано, что для критерия минимизации J_2 данный метод дает лучшее решение чем предложенные ранее алгоритмы, решающие наиболее близкую задачу и которыми можно было бы решать поставленную задачу. Для случая минимизации J_∞, решение предыдущих алгоритмов оказалось близко к оптимальному, но предложенный алгоритм все еще предпочтителен, т.к. гарантирует оптимальность решения.

Приложение:

Коэффициенты фильтров с параметрами $N = 10, M = 4$ в формате двумерного массива.

Минимизация $\boldsymbol{J_2}$, метод **CFG_POLY**: [[2.010e-3, 3.870e-4, -8.672e-3, -1.633e-3], [-1.336e-2, -3.503e-3, 5.748e-2, 1.476e-2], [5.065e-2, 1.931e-2, -2.163e-1, -8.089e-2], [-1.538e-1, -1.005e-1, 6.382e-1, 4.126e-1], [6.145e-1, 1.226e0, -4.708e-1, -9.219e-1], [6.145e-1, -1.226e0, -4.708e-1, 9.219e-1], [-1.538e-1, 1.005e-1, 6.382e-1, -4.126e-1], [5.065e-2, -1.931e-2, -2.163e-1, 8.089e-2], [-1.336e-2, 3.503e-3, 5.748e-2, -1.476e-2], [2.010e-3, -3.870e-4, -8.672e-3, 1.633e-3]].

Минимизация $\boldsymbol{J_2}$, метод **CFG_EXP**: [[-8.665e-3, 1.068e-2, -1.265e-2, 4.505e-3], [5.739e-2, -7.079e-2, 1.143e-1, -4.070e-2], [-2.155e-1, 2.663e-1, -6.263e-1, 2.231e-1], [6.309e-1, -7.850e-1, 3.189e0, -1.138e0], [9.981e-2, 5.150e-1, -4.624e0, 2.448e0], [9.981e-2, 5.150e-1, 4.624e0, -2.448e0], [6.309e-1, -7.850e-1, -3.189e0, 1.138e0], [-2.155e-1, 2.663e-1, 6.263e-1, -2.231e-1], [5.739e-2, -7.079e-2, -1.143e-1, 4.070e-2], [-8.665e-3, 1.068e-2, 1.265e-2, -4.505e-3]].

Минимизация $\boldsymbol{J_2}$, метод **CONT_POLY**: [[2.031e-3, 3.851e-4, -8.477e-3, -1.589e-3], [-1.342e-2, -3.473e-3, 5.589e-2, 1.426e-2], [5.066e-2, 1.909e-2, -2.101e-1, -7.794e-2], [-1.537e-1, -9.982e-2, 6.271e-1, 4.034e-1], [6.144e-1, 1.225e0, -4.646e-1, -9.062e-1], [6.144e-1, -1.225e0, -4.646e-1, 9.062e-1], [-1.537e-1, 9.982e-2, 6.271e-1, -4.034e-1], [5.066e-2, -1.909e-2, -2.101e-1, 7.794e-2], [-1.342e-2, 3.473e-3, 5.589e-2, -1.426e-2], [2.031e-3, -3.851e-4, -8.477e-3, 1.589e-3]].

Минимизация $\boldsymbol{J_2}$, метод **CONT_EXP**: [[-8.385e-3, 1.042e-2, -1.215e-2, 4.336e-3], [5.525e-2, -6.872e-2, 1.091e-1, -3.893e-2], [-2.074e-1, 2.582e-1, -5.962e-1, 2.128e-1], [6.163e-1, -7.706e-1, 3.087e0, -1.103e0], [1.087e-1, 5.062e-1, -4.447e0, 2.384e0], [1.087e-1, 5.062e-1, 4.447e0, -2.384e0], [6.163e-1, -7.706e-1, -3.087e0, 1.103e0], [-2.074e-1, 2.582e-1, 5.962e-1, -2.128e-1], [5.525e-2, -6.872e-2, -1.091e-1, 3.893e-2], [-8.385e-3, 1.042e-2, 1.215e-2, -4.336e-3]].

Минимизация $\boldsymbol{J_\infty}$, метод **CFG_POLY**: [[4.994e-5, 5.421e-3, 1.608e-2, 1.078e-1], [-7.231e-3, -2.791e-2, -1.913e-2, -4.047e-1], [4.087e-2, 8.247e-2, -9.021e-2, 8.954e-1], [-1.440e-1, -2.127e-1, 5.126e-1, -1.208e0], [6.104e-1, 1.373e0, -4.177e-1, 1.135e0], [6.104e-1, -1.373e0, -4.177e-1, -1.135e0], [-1.440e-1, 2.127e-1, 5.126e-1, 1.208e0], [4.087e-2, -8.247e-2, -9.021e-2, -8.954e-1], [-7.231e-3, 2.791e-2, -1.913e-2, 4.047e-1], [4.994e-5, -5.421e-3, 1.608e-2, -1.078e-1].

Минимизация $\boldsymbol{J_\infty}$, метод **CFG_EXP**: [[-5.970e-3, 8.197e-3, 1.324e-1, -4.558e-2], [3.908e-2, -5.313e-2, -4.795e-1, 1.642e-1], [-1.420e-1, 1.939e-1, 9.225e-1, -3.108e-1], [4.067e-1, -5.621e-1, -2.894e-1, 5.908e-2], [2.674e-1, 3.480e-1, 6.926e-1, 4.995e-1], [2.674e-1, 3.480e-1, -6.926e-1, -4.995e-1], [4.067e-1, -5.621e-1, 2.894e-1, -5.908e-2], [-1.420e-1, 1.939e-1, -9.225e-1, 3.108e-1], [3.908e-2, -5.313e-2, 4.795e-1, -1.642e-1], [-5.970e-3, 8.197e-3, -1.324e-1, 4.558e-2]].

Минимизация $\boldsymbol{J_\infty}$, метод **CONT_POLY**: [[1.342e-3, -2.086e-2, 3.529e-3, 2.777e-1], [-1.002e-2, 7.205e-2, 8.254e-3, -1.065e0], [4.511e-2, -1.455e-1, -1.320e-1, 2.407e0], [-1.475e-1, 1.642e-1, 5.457e-1, -3.715e0], [6.118e-1, 8.930e-1, -4.304e-1, 4.342e0], [6.118e-1, -8.930e-1, -4.304e-1, -4.342e0], [-1.475e-1, -1.642e-1, 5.457e-1, 3.715e0], [4.511e-2, 1.455e-1, -1.320e-1, -2.407e0], [-1.002e-2, -7.205e-2, 8.254e-3, 1.065e0], [1.342e-3, 2.086e-2, 3.529e-3, -2.777e-1]].

Минимизация $\boldsymbol{J_\infty}$, метод **CONT_EXP**: [[-8.527e-3, 1.064e-2, 1.326e-1, -4.670e-2], [5.112e-2, -6.470e-2, -5.096e-1, 1.791e-1], [-1.855e-1, 2.366e-1, 1.026e0, -3.579e-1], [5.425e-1, -6.969e-1, -2.856e-2, -1.019e-2], [1.647e-1, 4.502e-1, -7.974e-2, 8.153e-1], [1.647e-1, 4.502e-1, 7.974e-2, -8.153e-1], [5.425e-1, -6.969e-1, 2.856e-2, 1.019e-2], [-1.855e-1, 2.366e-1, -1.026e0, 3.579e-1], [5.112e-2, -6.470e-2, 5.096e-1, -1.791e-1], [-8.527e-3, 1.064e-2, -1.326e-1, 4.670e-2]].

Литература

1. C. W. Farrow, "A continuously variable digital delay element," *1988., IEEE International Symposium on Circuits and Systems*, Espoo, Finland, 1988, pp. 2641-2645 vol.3, doi: 10.1109/ISCAS.1988.15483.
2. Zhao, R., Tay, D.B. A Complex Exponential Structure for Low-Complexity Variable Fractional Delay FIR Filters. *Circuits Syst Signal Process* **42**, 1105–1141 (2023).
3. C. -C. Tseng and S. -L. Lee, "Design of fractional delay FIR filter using radial basis function," *2009 IEEE International Symposium on Circuits and Systems*, Taipei, Taiwan, 2009, pp. 485-488, doi: 10.1109/ISCAS.2009.5117791.
4. Chien-Cheng Tseng, "Design of sinusoid-based variable fractional delay FIR filter using weighted least squares method," *2003 IEEE International Symposium on Circuits and Systems (ISCAS)*, Bangkok, Thailand, 2003, pp. III-III, doi: 10.1109/ISCAS.2003.1205077.
5. Wu-Sheng Lu and Tian-Bo Deng, "An improved weighted least-squares design for variable fractional delay FIR filters," in *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, vol. 46, no. 8, pp. 1035-1040, Aug. 1999, doi: 10.1109/82.782046.
6. H. H. Dam, "Variable Fractional Delay FIR Filter Design with a Bicriteria and Coefficient Relationship," in *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 61, no. 1, pp. 36-40, Jan. 2014, doi: 10.1109/TCSII.2013.2291063.
7. R. Zhao and X. Hong, "Matrix-Based Algorithms for the Optimal Design of Variable Fractional Delay FIR Filters," in *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 66, no. 12, pp. 4740-4752, Dec. 2019, doi: 10.1109/TCSI.2019.2934140.
8. F. M. Gardner, "Interpolation in digital modems. I. Fundamentals," in *IEEE Transactions on Communications*, vol. 41, no. 3, pp. 501-507, March 1993, doi: 10.1109/26.221081.
9. Vesma, J., Saramaki, T. Polynomial-Based Interpolation Filters—Part I: Filter Synthesis. *Circuits Syst Signal Process* **26**, 115–146 (2007). https://doi.org/10.1007/s00034-005-0704-8
10. D. Babic, "Design of polynomial-based digital interpolation filters based on Chebyshev polynomials," *2013 36th International Conference on Telecommunications and Signal Processing (TSP)*, Rome, Italy, 2013, pp. 677-680, doi: 10.1109/TSP.2013.6614022.
11. J. Vesma, "A frequency-domain approach to polynomial-based interpolation and the Farrow structure," in *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, vol. 47, no. 3, pp. 206-209, March 2000, doi: 10.1109/82.826746.
12. Y. C. Lim, "Design of discrete-coefficient-value linear phase FIR filters with optimum normalized peak ripple magnitude," in *IEEE Transactions on Circuits and Systems*, vol. 37, no. 12, pp. 1480-1486, Dec. 1990, doi: 10.1109/31.101268.

OPTIMIZATION OF FARROW FILTER FOR RECONFIGURABLE INTERPOLATION PROBLEM

Busse A.A.[1], Solovev D.M.[1], Meshcheriakov A.A.[1]

[1] Moscow Institute of Physics and Technology

In this article new methodology is proposed for Farrow filter coefficients optimization using new criterion for reconfigurable interpolator problem. Currently existing methods consider either VFD (variable fractional delay) problem or targeting continuous signal reconstruction. Practically problems are often declared as interpolation with rational interpolation ratio, where interpolation ratio may differ from configuration to configuration, but all configurations are known in advance. This work motivation therefore to develop algorithm of Farrow filter optimization for such a problem to improve quality of existing solutions or prove they are close to the optimal. Problem is analyzed and optimal algorithms for two different criterions are proposed. Also, existing researches show that basis functions other than polynomial may provide better quality solutions. That's why proposed algorithms are generalized and may be used with any basis functions. For comparison basis consisting from complex exponents was implemented.

ОБРАБОТКА И АНАЛИЗ OFDM-СИГНАЛА С ШУМОМ НА ОСНОВЕ ВЕЙВЛЕТ-ПРЕОБРАЗОВАНИЯ

д.т.н., проф. Егоров В.В.[1], к.т.н., доц. Клионский Д.М.[2]

[1]Санкт-Петербургский государственный университет аэрокосмического приборостроения
[2]Санкт-Петербургский государственный электротехнический университет "ЛЭТИ"

Вопросы временной синхронизации являются актуальными для систем радиосвязи, радионавигации, радиолокации, определения моментов прихода импульсных сигналов, позиционирования и др. Так, в задаче радиосвязи необходимо обеспечить безошибочную передачу с передающего конца на приемный информации по радиоканалу с максимально возможной скоростью передачи. При этом точность синхронизации повышает скорость передачи информации. В известных решениях задачи синхронизации при передаче OFDM-сигналов широко применяется защитный интервал, что приводит к непроизводительным затратам временного ресурса. В предлагаемом в статье решении появляется возможность отказаться от использования защитного интервала и тем самым значительно повысить скорость передачи информации. В основе временной синхронизации заложен быстрый вычислительный алгоритм на основе гармонического вейвлет-преобразования, который может работать в режиме реального времени и является робастным по отношению к канальному шуму. В статье рассматривается задача обработки и анализа OFDM-сигналов в условиях шума и оценивания момента первого изменения параметров модели OFDM-сигнала.

Введение

В настоящее время в ряде практических задач в области телекоммуникаций часто применяются OFDM-сигналы (Orthogonal frequency-division multiplexing – мультиплексирование с ортогональным частотным разделением каналов) для высокоскоростной передачи данных по нестационарным радиоканалам, в т.ч. КВ-радиоканалам (коротковолновым радиоканалам) [1-2]. При обработке OFDM-сигналов принято использовать методы быстрых дискретных преобразований, что позволяет обрабатывать эти сигналы в реальном масштабе времени с использованием ортогонального частотного разделения субканалов. OFDM-сигналы используются в таких задачах, как частотная синхронизация, временная синхронизация, вычисление быстрых дискретных преобразований и т.д.

Длительность OFDM-сигнала и величина частотного разнесения выбираются из условия ортогональности сигналов на соседних субчастотах, при этом будет иметь место перекрытие спектров на соседних субчастотах.

Аналитическое описание (математическая модель) информационного OFDM-сигнала $s(t)$ имеет следующий вид [3]:

$$s(t)=\sum_{l=1}^{L}\sum_{k=1}^{K} c_k^{(l)} \cos\left(\omega_k\left(t-\Delta t_k\right)+\psi_k^{(l)}\right), \qquad (1)$$

где l - номер интервала ортогональности информационного OFDM-сигнала (обработка информационного сигнала на приемной стороне осуществляется по интервалам заданной длительности T), $l=\overline{1,\,L}$, L - общее число интервалов длительностью T излучаемого сигнала, K - общее количество субчастот многочастотного OFDM-сигнала, на которых происходит излучение, ω_k - k-я субчастота излучения, Δt_k - временная задержка k-го луча (k-й гармоники на субчастоте сигнала, $k=\overline{1,\,K}$) в канале, $c_k^{(l)}$ - амплитуда гармоники на k-й субчастоте на интервале ортогональности с номером l. Сигнал излучается интервалами длительностью T, при этом все частотные составляющие сигнала на частотах ω_k излучаются одновременно. Параметры $\left\{c_k^{(l)}\right\}$ и $\left\{\psi_k^{(l)}\right\}$ соответствуют k-й субчастоте излучения и l-му интервалу ортогональности сигнала на приемной стороне. Начальные фазы $\left\{\psi_k^{(l)}\right\}$, соответствующие l-му интервалу ортогональности, являются случайными величинами, зависящими от передаваемого символа, и изменяются при переходе от одной гармоники к другой.

Частота *k*-й гармоники излучения кратна периоду дискретизации по частоте и определяется следующими соотношениями:

$$\omega_k = k\Delta\omega,$$

$$\Delta\omega = \frac{2\pi}{T}.$$

На приемной стороне все частоты f_k претерпевают смещение δf:

$$\overline{f}_k = f_k + \delta f,\ k = \overline{1,K},$$

Частотное смещение происходит за счет допплеровского эффекта, а также за счет рассогласования частот принимаемого сигнала и опорного генератора.

На каждом интервале ортогональности математическая модель сигнала сохраняется неизменной, в соответствии с соотношением (1), однако при переходе от l-го к $l+1$-му интервалу ортогональности меняются параметры модели – величины $\left\{c_k^{(l)}\right\}$ и $\left\{\psi_k^{(l)}\right\}$ в выражении (1).

Структурная диаграмма информационного OFDM-сигнала.

OFDM-сигнал может быть представлен схематично в виде диаграммы, приведенной на рис. 1. На каждом интервале ортогональности длительностью T сигнал является мультигармоническим (полигармоническим), т.е. представляет собой конечный набор дискретных гармоник с амплитудами $\left\{c_k^{(l)}\right\}$, если руководствоваться представлением OFDM-сигнала в косинусной форме [3]. Интервал ортогональности также включает в себя защитный интервал, показанный на рис. 1. Защитный интервал длительностью $T_з$ представляет собой периодическое продолжение информационного сигнала и позволяет избавиться от влияния межсимвольных искажений при демодуляции. Межсимвольные помехи также возникают из-за нарушения ортогональности на соседних интервалах, если временная синхронизация была установлена неточно, что является критичным при использовании OFDM-сигналов.

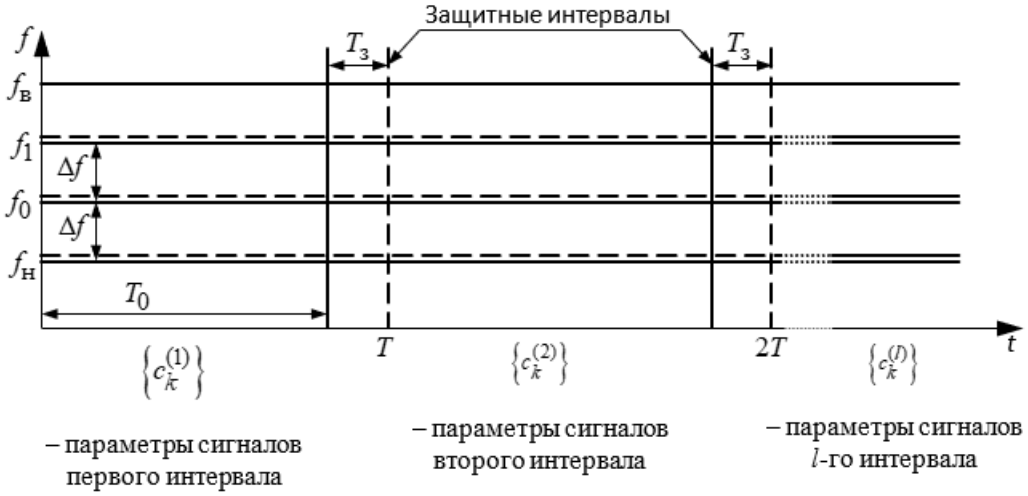


Рисунок 1 – Частотно-временная диаграмма информационного OFDM-сигнала

На рис. 1 несущие частоты отдельных каналов, составляющих OFDM-сигнал, расположены эквидистантно в диапазоне от $f_н$ до $f_в$. На приемной стороне дискретные составляющие на субчастотах претерпевают смещение, которое отображается на рисунке горизонтальными штриховыми линиями (причины возникновения частотного смещения описаны ранее).

Исследование OFDM-сигнала на приемной стороне в условиях шума.

Рассмотрим прием OFDM-сигнала в условиях белого гауссовского шума [4], что описано далее в вычислительном эксперименте. Определение временных границ базируется на использовании особых базисных функций с компактным носителем, называемых вейвлетами [5-6]. Носителем функции называют интервал, на котором функция принимает отличные от нуля зна-

чения, термин "компактный" подразумевает, что величина данного интервала (существенно) мала по сравнению с диапазоном задания функции [5-6]. На рис. 2 обозначены временные границы, соответствующие интервалам ортогональности OFDM-сигнала, а также оценки временных границ, полученные путем применения октавного гармонического вейвлет-преобразования [5-6].

Моменты времени t_l, соответствующие смене математической модели OFDM-сигнала, подчиняются соотношению:

$$t_l = lT + t_0,$$

где t_0 - момент первого изменения параметров модели OFDM-сигнала после начала анализа (наблюдения). Значение t_0 должно быть определено в процессе временной синхронизации.

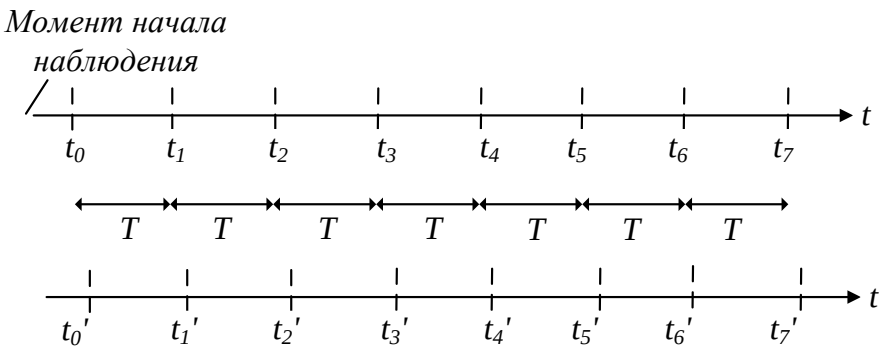


Рисунок 2 – Границы интервалов ортогональности OFDM-сигнала (вверху) и их оценки (внизу)

В первом случае (рис. 2, вверху) расстояние между моментами времени t_i и t_{i+1} для $i \geq 1$ одинаково и равно величине интервала ортогональности T. Во втором случае (рис. 2, внизу) границы, в общем случае, смещены относительно истинных значений ввиду наличия шума в анализируемом OFDM-сигнале. При различных отношениях сигнал/шум SNR определена дисперсия $\sigma^2_{уст}$ установления границ при временной синхронизации:

$$D = \sigma^2_{\text{уст}} = \frac{1}{L-1}\sum_{l=1}^{L}\left(t_l - t_l'\right)^2,$$

где t_l – истинное положение временно́й границы интервала ортогональности OFDM-сигнала, t_l' – оценка временно́й границы, полученная на основе гармонического вейвлет-преобразования, L – количество экспериментально определенных временны́х границ OFDM-сигнала (интервалов ортогональности).

Рассмотрим зашумленный OFDM-сигнал, включающий в себя 4 интервала ортогональности и 4 дискретных гармоники на каждом интервале. Среднеквадратическое отклонение шума составляет 0.1. Длина зашумленной реализации на одном интервале ортогональности OFDM-сигнала равна 4096 отсчетам (длина выбрана равной степени 2 для более точного применения гармонического вейвлет-преобразования) при периоде дискретизации 0.0001 с. Таким образом, длина всей временной реализации OFDM-сигнала составляет 16384 отсчетов. Шаг по частоте $\Delta\omega$ равен 15.3436 рад/с.

Параметры OFDM-сигнала, в соответствии с математической моделью (1), следующие:

- 1-й интервал ортогональности: $c_1^{(1)} = 0.8$, $c_2^{(1)} = 0.6$, $c_3^{(1)} = 0.4$, $c_4^{(1)} = 0.8$; $\psi_1^{(1)} = \pi/7$, $\psi_2^{(1)} = \pi/3$, $\psi_3^{(1)} = \pi/4$, $\psi_4^{(1)} = \pi/9$.
- 2-й интервал ортогональности: $c_1^{(2)} = 1.7$, $c_2^{(2)} = 0.4$, $c_3^{(2)} = 0.8$, $c_4^{(2)} = 0.15$; $\psi_1^{(2)} = \pi/2$, $\psi_2^{(2)} = \pi/9$, $\psi_3^{(2)} = 13\pi/12$, $\psi_4^{(2)} = \pi/3.5$.

- 3-й интервал ортогональности: $c_1^{(3)} = 0.9$, $c_2^{(3)} = 1.8$, $c_3^{(3)} = 0.1$, $c_4^{(3)} = 2.30$; $\psi_1^{(3)} = 3\pi/5$, $\psi_2^{(3)} = \pi/2$, $\psi_3^{(3)} = 8\pi/7$, $\psi_4^{(3)} = \pi/1.5$.
- 4-й интервал ортогональности: $c_1^{(4)} = 0.25$, $c_2^{(4)} = 0.15$, $c_3^{(4)} = 2.36$, $c_4^{(4)} = 1.12$; $\psi_1^{(4)} = -6\pi/5$, $\psi_2^{(4)} = -\pi/4$, $\psi_3^{(4)} = 2\pi/3$, $\psi_4^{(4)} = 0.2\pi$.

Частоты гармоник на всех 4-х интервалах ортогональности соответственно одинаковые и равны $\omega_1 = 2\Delta\omega$, $\omega_2 = 5\Delta\omega$, $\omega_3 = 4\Delta\omega$, $\omega_4 = 3\Delta\omega$.

Путем применения гармонического вейвлет-преобразования были установлены границы интервалов ортогональности, что обеспечило максимально возможную точность временной синхронизации.

На рис. 3 показана графическая зависимость дисперсии ошибки установления временного синхронизма от отношения сигнал/шум при различных длинах M_1 и M_2 окна анализа OFDM-сигнала.

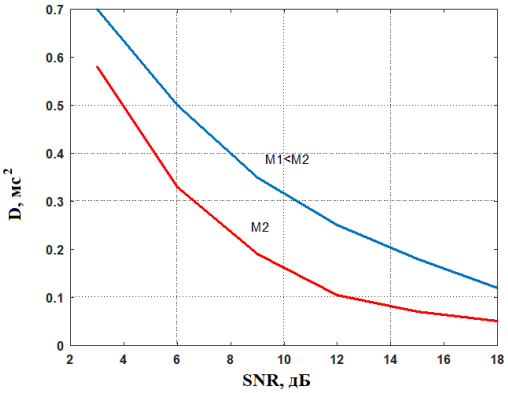


Рисунок 3 – Зависимость дисперсии ошибки установления временной синхронизации от отношения сигнал/шум при разных длинах окна анализа M_1 и M_2

Из данных зависимостей можно сделать вывод, что с увеличением окна анализа удается добиться более точного временного синхронизма (дисперсия определения временных границ интервалов ортогональности меньше), однако временной синхронизм при этом устанавливается дольше.

Установление временного синхронизма в условиях шума.

Представим совокупность временны́х границ интервалов ортогональности в виде массива $(t_1, t_2, ..., t_l, ..., t_L)$. Значения элементов массива $\{t_l\}_{l=1}^{L}$ необходимо определить, тем самым будет установлен временно́й синхронизм на приемной стороне.

В том случае если в исходном сигнале отсутствует шум, границы интервалов ортогональности OFDM-сигнала проецируются на единичную окружность в одну точку с единичной амплитудой, что можно аналитически представить в следующем виде:

$$e^{\frac{i2\pi}{T}t_l} = \{t_l = lT\} = e^{\frac{i2\pi}{T}lT} = e^{i2\pi l} = 1.$$

Таким образом, на комплексной плоскости имеем вектор в пределах единичной окружности, направленный по положительной оси.

При первом изменении параметров модели OFDM-сигнала после начала анализа в момент времени $t = t_0$ имеем следующее отображение:

$$e^{i(2\pi/T)t_l}\Big|_{\{t_l = t_0 + lT\}} = e^{i(2\pi/T)(lT+t_0)} = e^{i2\pi l}e^{i(2\pi/T)t_0} = e^{i(2\pi/T)t_0} = z_1.$$

Таким образом, в данном случае результат также инвариантен относительно l, имеет единичную амплитуду (т.е. расположен на единичной окружности), однако фаза $\arg\dot{Z}$ будет не нулевой, а равной $(2\pi/T)t_0$ и представит угол между вектором $\dot{Z}$ и осью абсцисс.

В реальных условиях в сигнале может присутствовать шум (например, шум канала передачи), в связи с чем оценки временны́х границ могут не полностью совпадать с их действительными положениями, т. е. быть не строго кратными значению интервала ортогональности. Кроме того, временны́е границы смещены на величину t_0, являющуюся моментом первого изменения параметров модели OFDM-сигнала после начала анализа. В результате отображение:

$$t_l \to e^{i(2\pi/T)t_l} = \{\dot{Z}_l\}, \tag{2}$$

представляет собой комплексный ансамбль. При наличии шума в OFDM-сигнале оценки временных границ $\hat{t}_l$ удовлетворяют соотношению:

$$\hat{t}_l = t_l + \varepsilon_l, \tag{3}$$

$$t_l = lT + t_0. \tag{4}$$

где t_l – истинные положения временны́х границ; ε_l – ошибка определения временно́й границы в силу наличия шума в исходном сигнале.

Далее вычисляется статистика $\dot{Z}_l$ по формуле (2):

$$e^{i(2\pi/T)\hat{t}_l} = \dot{Z}_l.$$

После этого суммируем значения статистики $\dot{Z}_l$ и получаем:

$$\dot{Z} = \sum_{l=1}^{L} \dot{Z}_l. \tag{5}$$

С учетом соотношений (3) и (4) выражение (5) можно представить в виде:

$$\begin{aligned} \dot{Z} &= \sum_{l=1}^{L} e^{i(2\pi/T)\hat{t}_l} = \sum_{l=1}^{L} e^{i(2\pi/T)(t_l+\varepsilon_l)} = \sum_{l=1}^{L} e^{i(2\pi/T)(t_0+lT+\varepsilon_l)} = \\ &= \sum_{l=1}^{L} e^{i(2\pi/T)t_0} e^{i2\pi l} e^{i(2\pi/T)\varepsilon_l} = e^{i(2\pi/T)t_0} \sum_{l=1}^{L} e^{i(2\pi/T)\varepsilon_l}. \end{aligned} \tag{6}$$

В выражении (6) присутствует постоянный множитель $e^{i(2\pi/T)t_0}$ и множитель $\sum_{l=1}^{L} e^{i(2\pi/T)\varepsilon_l}$, представляющий собой сумму комплексных экспонент с величинами ошибок ε_l в показателе степени. Из (6) получим соотношение

$$t_0 = \frac{\arg\dot{Z}}{2\pi} T.$$

определяющее момент первого изменения параметров модели OFDM-сигнала после начала анализа.

Векторная диаграмма, соответствующая OFDM-сигналу с шумом, представлена на рис. 4.

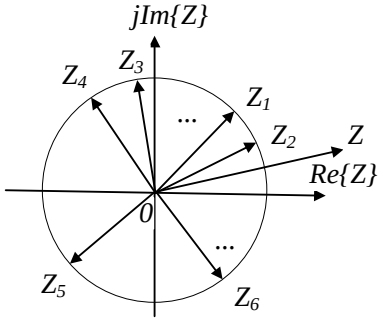


Рисунок 4 – Векторная диаграмма при наличии шума в OFDM-сигнале

На диаграмме показаны комплексные вектора $\dot{Z}_l$, соответствующие моментам времени $\hat{t}_l$, и результирующий комплексный вектор $\dot{Z}$, полученный суммой векторов $\dot{Z}_l$. Соответственно, $\arg \dot{Z}$ определяется как угол между вектором $\dot{Z}$ и осью абсцисс.

Заключение.

Октавное гармоническое вейвлет-преобразование является эффективным в задаче анализа и обработки OFDM-сигнала как при наличии, так и при отсутствии шума. Оно позволяет с максимально возможной точностью определять границы интервалов ортогональности OFDM-сигналов. На основе анализа комплексных векторов, соответствующих границам интервалов ортогональности, удается определить момент первого изменения параметров модели OFDM-сигнала в пределах окна анализа. Также получена зависимость дисперсии ошибки установления временного синхронизма от длины окна анализа, показывающая, что с увеличением окна анализа удается добиться более точного синхронизма (при этом временной синхронизм устанавливается дольше).

Литература

1) Бакулин М.Г., Крейнделин В.Б., Шлома А.М., Шумов А.П. Технология OFDM // учебное пособие для вузов, – М. : Горячая линия – Телеком, 2021, 360 с.

2) В.В. Егоров, А.Е. Тимофеев Установление частотно-временной синхронизации в многочастотных КВ-системах передачи данных // Электросвязь, № 7, 2013. – С. 41-44.

3) В.В. Егоров, Д.М. Клионский Применение гармонического вейвлет-преобразования при обработке OFDM-сигналов в нестационарном радиоканале // Цифровая обработка сигналов. 2024. № 2. С. 57-63.

4) Солонина А.И., Клионский Д.М., Меркучева Т.В., Перов С.Н. Цифровая обработка сигналов и MATLAB – Санкт-Петербург: БХВ-Петербург, 2013. – 512 с.

5) S. Mallat A wavelet tour of signal processing. Academic Press; First Edition, 1998, 577 p.

6) Н. К. Смоленцев Вейвлет-анализ в Matlab / – 3-е изд. – М. : ДМК Пресс, 2010. – 448 с.

OFDM SIGNAL ANALYSIS AND PROCESSING IN THE PRESENCE OF NOISE BASED ON THE WAVELET TRANSFORM

Dr. of Sci., Prof. V.V. Egorov[1], PhD, Ass. Prof. D.M. Klionskiy[2]

[1]Saint Petersburg State University of Aerospace Instrumentation
[2]Saint Petersburg Electrotechnical University "LETI"

The issues related to temporal synchronization are of great importance for various radio communication, radio navigation and radar systems. Temporal synchronization is employed when one needs to determine the signal arrival time and in the positioning problem. In radiocommunication, we need to provide accurate information transmission via a radio channel with a maximum possible transmission rate. Synchronization accuracy increases the information transmission rate. The well-known approaches to the synchronization problem in the case of OFDM signal transmission are based on using the guard interval, which leads to inefficiency in the time resource. The present paper suggests a new approach, which does not require guard interval use and, thereby, increases the information transmission rate. Temporal synchronization relies on using a special fast computational algorithm based on the harmonic wavelet transform capable of working in real time and being robustly resistant to channel noise. The paper also discusses some problems of OFDM signal processing and analysis in the presence of noise and estimation of the first change of the model parameters in an OFDM signal.

МАТЕМАТИЧЕСКИЙ МЕТОД НЕСЛЕПОГО ВОССТАНОВЛЕНИЯ СИГНАЛОВ С ПОМОЩЬЮ ТЕСТОВЫХ ИСПЫТАНИЙ СИСТЕМ ОБРАБОТКИ: ПОВЫШЕНИЕ ТОЧНОСТИ И ЭФФЕКТИВНОСТИ РЕКОНСТРУКЦИИ

с.н.с., к.т.н. Новиков-Бородин А.В.
Институт ядерных исследований РАН, 117312 Москва, Россия

Аннотация: Предлагается математический метод неслепого восстановления одномерных и многомерных сигналов, в том числе изображений, искаженных при обработке линейными стационарными системами. Вместо передаточных функций, которые зачастую трудно определить, т.к. они относятся к классу обобщенных функций, в предлагаемом методе для неслепого восстановления сигнала используются тестовые сигналы систем обработки. Использование тестовых сигналов, относящихся к классу основных функций, значительно упрощает процедуру восстановления сигналов и делает ее более точной и устойчивой. Операторный подход, основанный на уравнении многомерной свертки, существенно повышает скорость численных вычислений. Для решения некорректно поставленных и плохо обусловленных задач используется техника регуляризации, позволяющая эффективно восстанавливать реальные недетерминированные сигналы с высоким уровнем шумов и неопределенностей. Предлагаемая методика позволяет повысить качество обработки сигналов и изображений без использования сложных и дорогостоящих аппаратных методов модификации оборудования.
Программные коды и наборы данных доступны по адресу: *https://github.com/novikov-borodin/data-rec*.
Ключевые слова: математическая реконструкция сигналов и изображений; тестовые испытания систем обработки; многомерное уравнение типа свертки; некорректно поставленные и плохо обусловленные задачи; деконволюция; техника регуляризации.

Введение

Реальные системы обработки несовершенны, что приводит к системным искажениям и случайным шумам в обрабатываемых ими сигналах. Аппаратные методы уменьшения искажений связаны с решением сложных научно-технических задач и созданием дорогостоящего оборудования. Альтернативные математические методы снижения системных искажений и шумов в сигналах сталкиваются с серьезными трудностями при решении некорректно поставленных и плохо обусловленных задач реконструкции (см., например, [1]).

Преобразования сигналов линейных стационарных систем, рассматриваемых в данной работе, удовлетворяют уравнениям типа свертки (см. раздел 1) и описываются передаточными или аппаратными функциями, характеризующими системные искажения сигналов. Неслепые методы математической реконструкции являются наиболее общим подходом к восстановлению сигналов, но требуют знания передаточных функций систем, определить которые достаточно трудно, т.к. они зачастую принадлежат к классу обобщенных функций (см., например, [2]), что резко ограничивает точность и область применения неслепых методов некоторыми частными случаями известных видов искажений и шумов (см., например, [3-5]).

Методы слепого восстановления (см., например, [6,7]) позволяют избежать проблем с определением передаточных функций, но они основаны на априорных знаниях или предположениях об исходном или желаемом виде реконструируемых данных. Реальные сигналы, искажения и шумы, вносимые системами обработки, очень разнообразны, что сильно ограничивает область применения подобных методов. Использование нейронных сетей (см., например, [8]) позволяет обойтись без прямой идентификации передаточных функций, но обучение нейросетей происходит на большом количестве примеров, что дает усредненные подходы и во многих случаях не обеспечивает высокой точности реконструкции. По общим оценкам, приведенным в пленарном докладе [9], существующие математические методы, включая нейронные сети, повышают эффективные параметры измерительных систем (например, эффективную полосу пропускания) в среднем в $1{,}5 - 2$ раза. При использовании предлагаемого в данной работе подхода эффективную полосу пропускания системы обработки (см. примеры в разделе 3) удалось улучшить на три порядка – с 1 МГц до ~1 ГГц.

Предлагаемый в данной работе метод является неслепым методом восстановления как одно-, так и многомерных сигналов, включая изображения, т.е. является наиболее общим подходом к устранению разнообразных системных искажений различных данных. В то же время этот метод не требует определения передаточных функций систем обработки, вместо которых используются тестовые сигналы этих систем. В отличие от передаточных функций, тестовые сигналы

принадлежат к классу основных функций, поэтому их определение, возможное как численными, так и экспериментальными методами, не вызывает больших проблем и определяет простоту, универсальность и точность предлагаемого подхода. Также при отсутствии передаточных функций из класса обобщенных появляется возможность эффективного использования существующих методов решения некорректно поставленных и плохо обусловленных задач реконструкции, таких как методы регуляризации (см., например, [10]), используемые в данной работе. Все это обеспечивает универсальность предлагаемого подхода, высокую устойчивость алгоритма и хорошую точность восстановления данных даже при высоком уровне шумов и неопределенностей в них.

1. Уравнение тестовых испытаний

Функциональная система обработки сигналов и изображений является линейной стационарной (ЛС), если ее работа в области $x \in D$ характеризуется передаточной функцией $u(x)$ и ее реакция $w(x)$ (выходной сигнал) на внешнее воздействие $v(x)$ (входной сигнал) описывается уравнением типа свертки (см., например, [2]), которое будем называть уравнением ЛС-систем, или, кратко, ЛС-уравнением:

$$w(x) - \{\eta(x)\} = (u * v)(x) = \int_D u(x-\xi)v(\xi)\mathrm{d}\xi. \qquad (1)$$

Здесь $\eta(x) = (\Delta w - u * \Delta v - \Delta u * v + \Delta u * \Delta v)(x)$ – неопределенность ЛС-уравнения, которая следует из уравнения свертки: $w_0(x) = (u_0 * v_0)(x)$ функций без неопределенностей: $w_0(x) = (w - \Delta w)(x)$, $u_0(x) = (u - \Delta u)(x)$ и $v_0(x) = (v - \Delta v)(x)$, показанных на рисунке 1.

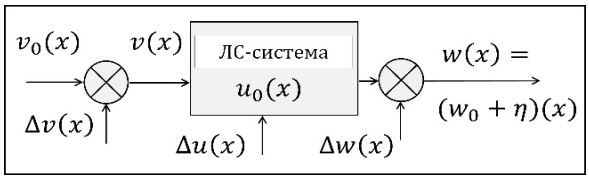


Рисунок 1 – Обработка сигналов ЛС системой.

Вводя оператор свертки $\mathcal{F}$ с порождающей функцией $f(x)$: $\mathcal{F}g(x) = (f * g)(x)$, можно представить ЛС-уравнение в операторной форме: $\mathcal{U}v(x) = w(x) - \{\eta(x)\}$, а используя теорему Бореля о свертке, – в спектральной форме: $U(\omega)V(\omega) = W(\omega) - \{N(\omega)\}$, где $F(\omega) = FT[f(x)]$ – спектр (преобразование Фурье) функции $f(x)$, а $N(\omega) = \Delta W(\omega) - U(\omega)\Delta V(\omega) - \Delta U(\omega)V(\omega) - \Delta U(\omega)\Delta V(\omega)$.

Помимо уравнения (1) ЛС-системы можно описать с помощью обыкновенных дифференциальных уравнений с постоянными коэффициентами a_k: $w(x) = \sum_k a_k v^{(k)}(x)$. При этом, так как для локально интегрируемых функций $v^{(k)}(x) = \left(\delta^{(k)} * v\right)(x)$, где $\delta^{(k)}(x)$ – k-тая производная дельта-функции, передаточная функция будет принадлежать к классу обобщенных: $u(x) = \sum_k a_k \delta^{(k)}(x)$, что создает серьезные проблемы при неслепой реконструкции сигналов.

Например, определение передаточной функции $u_0(x)$ из класса обобщенных функций из уравнения тестовых сигналов (ТС-уравнения):

$$(u_c * v_c)(x) = w_c(x) - \{\eta_c(x)\}, \qquad (2)$$

где $w_c(x)$ – реакция ЛС-системы на некоторое тестовое воздействие $v_c(x)$, а $\eta_c(x) = (\Delta w_c - u_c * \Delta v_c - \Delta u_c * v_c - \Delta u_c * \Delta v_c)(x)$ и $u_c(x) = (u_0 + \Delta u_c)(x)$ (см. Рис.1), может оказаться гораздо более трудной задачей, чем восстановление из уравнения (1) сигналов, относящихся к классу основных функций.

Многих проблем, связанных с использованием передаточных функций, можно избежать, исключив их из рассмотрения. Свернем обе части ТС-уравнения (2) с воздействием $v(x)$, а ЛС-уравнения (1) с тестовым воздействием $v_c(x)$:

$$\begin{cases} (v * w_c)(x) = \left(v * (u_c * v_c)\right)(x) + \{(v * \eta_c)(x)\}, \\ (v_c * w)(x) = \left(v_c * (u * v)\right)(x) + \{(v_c * \eta)(x)\}. \end{cases}$$

Вычитая второе уравнение из первого: $(v * w_c - v_c * w)(x) = \left(v * (u_c * v_c) - v_c * (u * v)\right)(x) + \{(v * \eta_c - v_c * \eta)(x)\}$ и учитывая, что: $(v * u_c * v_c - v_c * u * v)(x) = \left(v * v_c * (\Delta u_c - \Delta u)\right)(x)$,

т.к. $(u_c - u)(x) = (u_0 + \Delta u_c - u_0 - \Delta u)(x) = (\Delta u_c - \Delta u)(x)$, получим уравнение тестовых испытаний (ТИ-уравнение):

$$(w_c * v)(x) = (v_c * w)(x) - \{\eta_T(x)\}. \qquad (3)$$

Неопределенность ТИ-уравнения: $\eta_T(x) = \big(v_c * (\eta + v * \Delta u) - v * (\eta_c + v_c * \Delta u_c)\big)(x)$ меньше, чем общая неопределенность ЛС- и ТС-уравнений: $\eta_T(x) \le (v_c * \eta - v * \eta_c)(x)$, т.к. члены $v * \Delta u$ и $v_c * \Delta u_c$, входящие в η и η_c с противоположным знаком, сокращаются. Следовательно, ожидаемая погрешность реконструкции сигналов из ТИ-уравнения будет меньше, чем из ЛС- и ТС-уравнений. Кроме того, ТИ-уравнение не содержит передаточных функций и при его использовании не возникает ошибок представления функций из класса обобщенных.

Если далее в выражении для $\eta_T(x)$ учесть, что $u(x) = (u_0 + \Delta u)(x)$ и $u_c(x) = (u_0 + \Delta u_c)(x)$, то можно видеть, что неопределенность ТИ-уравнения $\eta_T(x) = \big(v_c * (\Delta w - u_0 * \Delta v) - v * (\Delta w_c - u_0 * \Delta v_c)\big)(x)$ не зависит явно от неопределенностей $\Delta u(x)$ и $\Delta u_c(x)$ передаточной функции, что повышает стабильность алгоритма восстановления сигналов из ТИ-уравнения.

Отметим, что определение сигналов ЛС-систем, входящих в ТИ-уравнение и относящихся к функциям из основного класса, не вызывает затруднений, поскольку могут быть использованы как численные, так и экспериментальные методы. Более того, при экспериментальном определении тестовых сигналов не требуется ни анализ ЛС-системы, ни учет паразитных параметров ее элементов, поскольку эта информация содержится в тестовых сигналах. Это позволяет восстанавливать сигналы высокоскоростных и прецизионных систем обработки практически любой сложности и структуры. Например, ТИ-методы можно использовать для улучшения параметров осциллографов. В отличие от работы [5], ТИ-методы не требуют определения передаточных функций и преобразований с ними, а тестовые сигналы непосредственно используются при реконструкции, что значительно упрощает процедуру реконструкции, и, согласно приведенному анализу, должно привести к повышению ее точности.

2. Решение ТИ-уравнений

ТИ-уравнение является уравнением типа свертки и, согласно теории регуляризации (см., напр., [20-22]), его решением $v_r(x)$ будет функция, минимизирующая функционал $\Phi[v] = \|\mathcal{V}_c w - \mathcal{W}_c v\|^2 + \lambda\|\mathcal{P}v\|^2$ со стабилизатором $\lambda\|\mathcal{P}v\|^2$, где $\mathcal{P}$ – оператор, соответствующий ограничениям, накладываемым на реконструируемую функцию $v(x)$, а λ – параметр регуляризации. Приравнивая к нулю вариацию этого функционала по v, получим:

$$v_r(x) = (\mathcal{W}_c^+ \mathcal{W}_c + \lambda \mathcal{P}^+ \mathcal{P})^{-1} \mathcal{W}_c^+ \mathcal{V}_c w(x). \qquad (4)$$

Здесь $\mathcal{F}^+$ – сопряженно-транспонированный оператор к оператору $\mathcal{F}$. В спектральной форме это решение имеет вид: $V_r(\omega) = W_c^+(\omega) V_c(\omega) W(\omega) / \big(W_c^+(\omega) W_c(\omega) + \lambda P^+(\omega) P(\omega)\big)$.

В многомерном случае равенство $FT(\mathcal{F}^+\mathcal{F}) = F^+(\omega)F(\omega)$ можно рассматривать как определение сопряженно-транспонированного оператора.

Для действительных функций со случайными некоррелированными неопределенностями, обычно встречающимися на практике, оптимальный стабилизатор соответствует фильтру Винера: $\lambda P^+(\omega)P(\omega) = |N_T(\omega)|^2 / |V(\omega)|^2$, где под спектрами мощности $|N_T(\omega)|^2$ и $|V(\omega)|^2$ случайных стационарных процессов понимаются ожидаемые значения. Стабилизатор Винера обычно неизвестен, но поскольку он неотрицателен и симметричен, его ищут в виде неотрицательной симметричной функции, например, в виде полинома: $\lambda P^+(\omega)P(\omega) = \sum_{k=0}^{n} \lambda_k \omega^{2k}$ с неотрицательными коэффициентами $\lambda_k \ge 0$, рассматриваемыми как параметры регуляризации. Согласно преобразованию Фурье для производных, это представление соответствует решениям, имеющим производные порядка $2k$ со “степенью непрерывности”, определяемой значениями λ_k. Следовательно, наиболее широкий класс решений со “степенью непрерывности” определяемой значением λ_0 дает стабилизатор нулевого порядка ($n = 0$): $P^+(\omega)P(\omega) = \lambda_0$ или $\mathcal{P}^+\mathcal{P} = \lambda_0 \mathcal{J}$, где $\mathcal{J}$ - тождественный оператор.

Остаток $\big\|\big(v_c * w - w_c * v_r(\lambda)\big)(x)\big\|$ монотонно возрастает с увеличением λ, поэтому оптимальные значения λ можно найти из уравнения: $\big\|\big(v_c * w - w_c * v_r(\lambda)\big)(x)\big\| = \|\eta_T(x)\|$. Отсюда следует необходимое условие для восстановления сигналов из ТИ-уравнения: $\|\eta_T(x)\| \le \|(v_c * w - w_c * v)(x)\|$.

Пусть функции $g(x)$, $f(x)$ и $h(x)$: $g(x) = (f * h)(x)$ представлены в виде M-мерных массивов $g = \{g_{i_1..i_M}\}$, $f = \{f_{i_1..i_M}\}$ и $h = \{h_{i_1..i_M}\}$ на сетке $\{d_{i_1....i_M}\}$, $i_k = 0 \dots m_k - 1, k = 1 \dots M$, разбивающей прямоугольную область $D_0 \supseteq D$. Для элемента $g_{i_1..i_M} = g\left(d_{i_1..i_M}\right)$ имеем: $g_{i_1..i_M} = \int_{D_0} f(\xi) h\left(d_{i_1..i_M} - \xi\right) \mathrm{d}\xi$. Частичная сумма этого интеграла на сетке $\{d_{i_1..i_M}\}$ представляет собой дискретное уравнение свертки:

$$g_{i_1..i_M} \approx \sum_{j_1=0}^{m_1-1} \cdots \sum_{j_M=0}^{m_M-1} f_{j_1..j_M} h_{i_1-j_1..i_M-j_M} \Delta_{j_1..j_M}, \qquad (5)$$

где массив $f_{j_1..j_M}$ переопределен как $f_{j_1..j_M} = f_{j_1..j_M}\Delta_{j_1..j_M}$, где $\Delta_{j_1..j_M}$ – элемент объема.

Определяя частичную сумму в уравнении (5) как оператор свертки многомерных дискретных данных: $g = \mathcal{F}h = f * h$, из уравнения (2) получим дискретное ТИ-уравнение: $\mathcal{V}_c w = \mathcal{W}_c v - \{\eta_T\}$. На равномерной сетке разбиения элементы объема в обеих частях этого уравнения сокращаются, поэтому, в отличие от уравнения (5), переопределение функций здесь не требуется. Из уравнения (4) получим дискретное регуляризированное решение ТИ-уравнения в операторном:

$$v_r = (\mathcal{W}_c^+ \mathcal{W}_c + \lambda \mathcal{P}^+ \mathcal{P})^{-1} \mathcal{W}_c^+ \mathcal{V}_c w \qquad (6)$$

и в спектральном виде: $V^r_{j_1..j_M} = W^{c+}_{j_1..j_M} V^c_{j_1..j_M} W_{j_1..j_M} / \left(W^{c+}_{j_1..j_M} W^c_{j_1..j_M} + \lambda P^+_{j_1..j_M} P_{j_1..j_M}\right)$.

Количество операций необходимых для восстановления массивов с n элементами с помощью операторного уравнения (6), основанного на уравнении дискретной свертки (5), может быть оценено как: $N(n) \sim \mathrm{O}\left(5n^2 + n + N(\lambda\mathcal{P})\right)$, что для больших n намного быстрее, чем матричный подход, использовавшийся ранее в работе [18] и требовавший $\sim 5n^3$ операций. Необходимое количество операций при использовании спектральной формы уравнения (6) (с прямым и обратным ДПФ) равно: $N(n) \sim \mathrm{O}\left(4n\log(n) + n + N(\lambda P)\right)$. Здесь $N(\lambda\mathcal{P})$ и $N(\lambda P)$ – количество операций со стабилизаторами.

Таким образом, спектральные уравнения (6) значительно быстрее операторных, но при численных расчетах на конечных областях D операторные уравнения дают нестационарные решения, а спектральные – установившиеся, с граничными условиями, соответствующими периодическим сигналам (см. раздел 1), что при, как правило, неопределенных граничных условиях резко ограничивает область применения спектральных методов.

3. Примеры реконструкции одномерных сигналов и двумерных изображений

Результаты численного эксперимента по восстановлению одномерных сигналов $v_0(t) = (v - \Delta v)(t)$ сложной формы (суперпозиция трапецеидального и гауссово импульсов) на входе емкостного делителя напряжения из искаженного выходного сигнала $w(t)$ по тестовым сигналам $w_c(t)$ и $v_c(t)$ представлены на рисунке 2. В качестве тестового воздействия $v_c(t)$ использовался трапецеидальный импульс с разными экспоненциальными фронтами, обеспечивающий широкий спектр сигнала (подробности см. в работе [11]). Принципиальная схема делителя была взята из стандартной библиотеки симулятора аналоговых/цифровых схем Micro-CAP 12.2.0.4 компании Spectrum Software. Исходные сигналы были рассчитаны также с помощью этой программы. Для наглядности выходные сигналы показаны с коэффициентом масштабирования, компенсирующим падение напряжения на делителе.

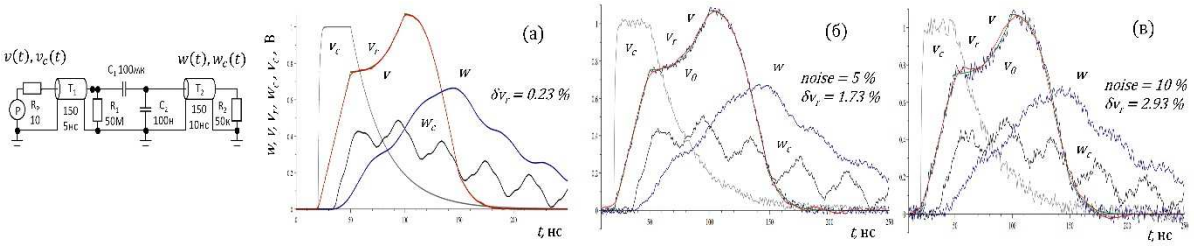


Рисунок 2 – Восстановление 1D сигналов

Все сигналы на интервале от 0 до 250 нс были преобразованы в векторы w, v, v_0, w_c, v_c размера $n = 500$ с шагом дискретизации $\Delta = 0.5$ нс. Для реконструкции использовалось

операторное уравнение (6) со стабилизатором нулевого порядка $\lambda_0 \mathcal{I}$, что соответствует поиску решений наиболее широкого класса со “степенью непрерывности”, определяемой параметром λ_0. Без дополнительных шумов в сигналах относительная ошибка реконструкции $\mathbf{v}_r$ сигнала $\mathbf{v}_0$ в точках выборки составила $\delta v_r = \|v_r - v_0\|/\|v_0\| = 0.23$ %. Результаты v_r восстановления сигнала v_0 при наличии дополнительного гауссовского шума с дисперсией 5 и 10 % от максимальных значений сигналов v, v_c и w, w_c показаны на рисунках 2б и 2в. Относительные ошибки реконструкции δv_r в этих случаях составили 1.73 и 2.93 %.

В рассмотренном примере частота среза делителя составляла 1 МГц, но с помощью предложенного математического метода даже при высоком уровне шумов в данных удалось восстановить сигналы с верхней границей спектра ~ 1 ГГц. Таким образом, эффективная полоса пропускания системы обработки была улучшена на три порядка.

Численные расчеты проводились в среде Maple 2017.3 Waterloo Maple Inc. на ноутбуке HP 255 G7 с процессором AMD Ryzen 3, 2.5 ГГц. Коды программ доступны в библиотеке GitHub по адресу *https://github.com/novikov-borodin/data-rec/signal-rec*. Время восстановления дискретных сигналов размером $n = 501$, соответствующих шагу дискретизации $\Delta = 0.5$ нс составило 1.6..2.0 сек. При размере векторов $n = 251$ ($\Delta = 1.0$ нс) время вычислений составило 0.4…0.5 сек, что соответствует сделанным ранее оценкам квадратичной зависимости количества операций (см. комментарии к уравнению (6) в разделе 2).

На рисунке 3 показан результат восстановления $\mathbf{v}_r$ цветного изображения $\mathbf{v}$ размером 500×700 пикселей в формате *jpeg* из изображения $\mathbf{w}$ размера 549×769, размытого с ядром $\mathbf{u}$ размера 50×70 с помощью спектрального представления уравнения (6).

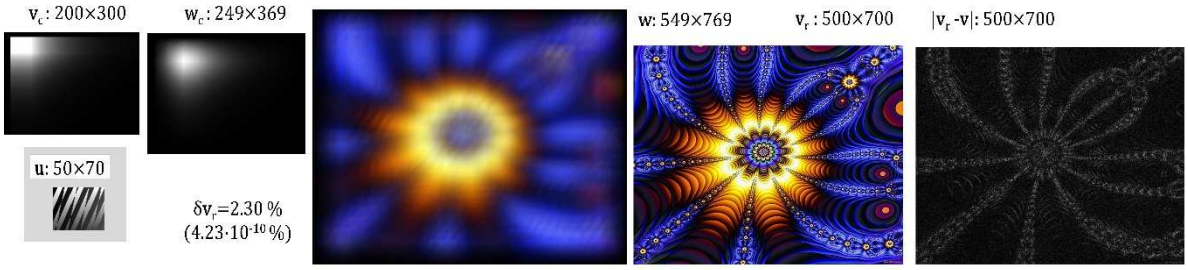


Рисунок 3 – Восстановление изображений

Расчеты проводились в среде Maple 2017.3 Waterloo Maple Inc. Коды программ доступны по адресу *https://github.com/novikov-borodin/data-rec/image-rec*. Время восстановления цветного (трехслойного) *jpeg* изображения размером 500×700 пикселей, показанного на Рис.3, составило 127.2 сек, что соответствует времени вычисления каждого слоя $t = 127.2/3 = 42.4$ сек. Время восстановления *jpeg* изображения в оттенках серого (с одним слоем) размером 201×201 пикселей составило $t_2 = 4.453$ сек. Пересчет времени восстановления $t = t_2 N(n)/N(n_2) = 42.55$ сек (≈ 42.4 сек) показал хорошее соответствие с приведенными в разделе 2 оценками необходимого количества операций.

Заключение

Предложены и проанализированы математические методы восстановления многомерных сигналов и изображений, искаженных и зашумленных при их обработке линейными стационарными системами, с помощью тестовых испытаний этих систем. В предлагаемых неслепых методах реконструкции вместо передаточных функций систем обработки, являющихся также функциями искажения сигналов, используются тестовые сигналы систем, что значительно упрощает процедуру восстановления сигналов и делает ее более устойчивой и точной.

Операторный подход, основанный на уравнении многомерной свертки, позволяет восстанавливать многомерные сигналы и изображения и существенно повышает скорость численных расчетов. Для решения некорректно поставленных и плохо обусловленных задач реконструкции используется метод регуляризации, позволяющий эффективно восстанавливать реальные недетерминированные сигналы с шумами и неопределенностями.

Предложенная методика расширяет возможности применения методов математической реконструкции на практике, позволяет повысить качество обработки сигналов и изображений без сложной и дорогостоящей модернизации оборудования. Методика имеет широкую область практического применения.

Литература

1. Kazufumi I., Bangti J. Inverse Problems: Tikhonov Theory and Algorithms // Singapore: World Scientific, 2014.
2. Sobolev V.I. Convolution of functions, Encyclopedia of Mathematics // EMS Press, Helsinki, 2001.
3. Chaganova O.B., Grigoryev A.S., Nikolaev D.P., Nikolaev I.P. Applied aspects of modern non-blind image deconvolution methods // Computer Optics, 48(4): 562-572, 2024. doi: 10.18287/2412-6179-CO-1409.
4. Wang H. et al. Ambulatory ECG noise reduction algorithm for conditional diffusion model based on multi-kernel convolutional transformer // The Review of scientific instruments. 95(9): 95-107, 2024.
5. Kleopin A.V., Zenchenko M.A. Tikhonov Regularization in Pulse Signal Processing for Oscilloscope Measurements // XXIV International Conference on Digital Signal Processing and its Applications (DSPA), 2022, pp. 1-4. doi: 10.1109/DSPA53304.2022.9790777.
6. Levin A., et al. Understanding and evaluating blind deconvolution algorithms // 2009 IEEE conference on computer vision and pattern recognition, IEEE, 2009.
7. Junjie T. A.O., et al. An image defocus deblurring method based on gradient difference of boundary neighborhood // Virtual Reality & Intelligent Hardware, 5(6): 538-549, 2023. doi: 10.1016/j.vrih.2023.06.008.
8. Hyeontae Jo, Hwijae Son, Hyung Ju Hwang, Eun Heui Kim. (2020). Deep neural network approach to forward-inverse problems. // Networks and Heterogeneous Media, 15(2): 247-259. doi: 10.3934/nhm.2020011.
9. Malay I. Digital Signal Processing in Modern Radio Engineering Measurements // Plenary Talk at XXV International Conference Digital Signal Processing and Its Applications – DSPA-2023, 2023.
10. Elsayed M.A., El-Shafai W., Rashwan M.A., Dessouky M.I., El-Fishawy A.S., Abd El-Samie F.E. Efficient iterative implementation of regularized solutions for image and signal reconstruction problems // Journal of Optics, 2023. doi: 10.1007/s12596-023-01179-w.
11. Novikov-Borodin A.V. Test Methods for Signal Reconstruction of Linear Stationary Systems // Journal of Communications Technology and Electronics, vol. 68(7): 732-747, 2023. doi: 10.1134/S1064226923070082.

Mathematical Method for Non-Blind Signal Recovery Using Test Trials of Processing Systems: Improving the Accuracy and Efficiency of Reconstruction

A.V. Novikov-Borodin
Institute for Nuclear Research of RAS, 117312 Moscow, Russia

Abstract. The mathematical method of non-blind recovery of one-dimensional and multidimensional signals, including images, distorted during processing by linear stationary systems is proposed. Instead of transfer functions, which are often difficult to determine because they belong to the class of generalized functions, the proposed method uses test signals of processing systems for non-blind signal recovery. The use of test signals belonging to the class of the main functions greatly simplifies the signal recovery procedure and makes it more accurate and stable. The operator approach based on the multidimensional convolution equation significantly increases the speed of numerical computations. To solve ill-posed and ill-conditioned problems, a regularization technique is used to efficiently recover real nondeterministic signals with noise and uncertainties. The proposed technique is able to improve the quality of signal and image processing without using complex and expensive hardware methods of equipment modification.
Program codes and datasets are available in GitHub library at: *https://github.com/novikov-borodin/data-rec*.
Keywords: mathematical reconstruction of signals and images; test trials of processing systems; multidimensional convolution type equation; ill-posed and ill-conditioned problems; deconvolution; regularization technique.

ИСПОЛЬЗОВАНИЕ ЛС-ПРЕОБРАЗОВАНИЙ ДЛЯ РЕКОНСТРУКЦИИ, КОРРЕКЦИИ И МОДЕЛИРОВАНИЯ ЭКСПЕРИМЕНТАЛЬНЫХ ДАННЫХ

с.н.с., к.т.н. Новиков-Бородин А.В.
Институт ядерных исследований РАН, 117312 Москва, Россия

Аннотация: Рассматриваются возможности применения математических методов ЛС-преобразований для обработки экспериментальных данных, искаженных линейными стационарными системами. Методы позволяют осуществлять коррекцию, реконструкцию и моделирование сигналов систем обработки в важных и часто встречающихся на практике случаях, когда большинство существующих как прямых, так и приближенных математических методов оказываются неэффективными. Рассматриваются примеры повышения эффективных параметров качества экспериментального оборудования, включая электронные приборы, устройства приёма-передачи, спектрометры и т.д., вплоть до сверх высокого разрешения. Предлагаемые методы позволяют повысить качество обработки данных без использования сложных и дорогостоящих аппаратных методов модификации оборудования.

Ключевые слова: математическая обработка сигналов и изображений; техника ЛС-преобразований; линейные стационарные системы; многомерное уравнение типа свертки.

Введение

Несовершенство систем обработки приводит к появлению искажений в получаемых, преобразуемых и передаваемых ими данных. Аппаратные методы уменьшения искажений связаны с решением сложных научно-технических и технологических задач, а также с созданием дорогостоящего оборудования. Альтернативой им являются математические методы реконструкции неискаженных данных из искаженных в процессе обработки (см., например, [1-3] и ссылки в них). Однако, реальные недетерминированные сигналы содержат шумы и неопределенности, что делает задачи реконструкции некорректно поставленными и плохо обусловленными. Универсального подхода к решению подобных задач не существует, а известные методики или достаточно сложны, или применимы далеко не всегда.

Предлагаемые в данной работе математические методы ЛС-преобразований (ЛСП-методы) предназначены для коррекции, реконструкции и моделирования сигналов и изображений, искаженных в результате их обработки линейными стационарными (ЛС) системами (см., например, [4,5]). ЛСП-методы применимы для обработки как дискретных, так и непрерывных сигналов, включая многомерные. Они относятся к методам прямой инверсии и практически не ограничивают класс реконструируемых функций, которые в рассматриваемой области должны быть интегрируемыми по Лебегу, что позволяет обрабатывать кусочно-непрерывные функции с разрывами первого рода, сигналы спектрометров с узкими спектральными линиями, контрастные изображения с резкими границами и т.п.

Прямые методы (см., например, [2,3]) неприемлемы для решения некорректно поставленных и плохо обусловленных задач, а приближенные методы, такие как техника регуляризации (см., например, [6,7]), не подходят для восстановления функций столь широкого класса. ЛСП-методы в отличие от других прямых методов (прямая инверсия, псевдоинверсия Мура-Пенроуза, матричная факторизация и др.) позволяют решить задачу с минимальными погрешностями в ряде важных и часто встречающихся на практике случаях за счет возможности выбора типа ЛС-преобразований и сокращения их числа. В работе представлены варианты решения базовых задач и примеры практического использования ЛС-преобразований.

1. ЛС-преобразования

Функциональная система обработки одно- и многомерных сигналов, включая изображения, является линейной стационарной, если ее работа в области $x \in D$ характеризуется передаточной функцией $u(x)$ и ее реакция $w(x)$ (выходной сигнал) на внешнее воздействие (x) (входной сигнал) описывается уравнением типа свертки (см., например, [5]), которое будем называть уравнением ЛС-систем, или, кратко, ЛС-уравнением:

$$w(x) - \{\eta(x)\} = (u * v)(x) = \int_D u(x-\xi)v(\xi)d\xi. \qquad (1)$$

Здесь $\eta(x)$ – неопределенность ЛС-уравнения, являющаяся следствием наличия шумов и неопределенностей в реальных недетерминированных функциях $w(x)$, $v(x)$ и $u(x)$. Вводя оператор свертки $\mathcal{S}$ с порождающей функцией $s(x)$ (в общем случае обобщенной): $\mathcal{S}f(x) = (s * f)(x)$, ЛС-уравнение можно представить в операторной форме: $\mathcal{U}v(x) = w(x) - \{\eta(x)\}$, или, используя теорему Бореля о свертке, – в спектральной форме: $W(\omega) - \{N(\omega)\} = U(\omega)V(\omega)$, где $S(\omega) = FT[\mathcal{S}] = FT[s(x)]$ – спектр (преобразование Фурье) функции $s(x)$ или оператора $\mathcal{S}$.

Согласно линейности, коммутативности и ассоциативности свертки последовательное применение $\mathcal{S} = \mathcal{S}_n \ldots \mathcal{S}_2\mathcal{S}_1$ и линейные комбинации $\mathcal{S} = \sum_i a_i\mathcal{S}_i$ операторов свертки, являющиеся операторами свертки, будем называть ЛС-преобразованиями. Осуществляя ЛС-преобразования $\mathcal{S}$ над обеими частями ЛС-уравнения и учитывая коммутативность и ассоциативность свертки, получим уравнение ЛС-преобразований (ЛСП-уравнение):

$$\mathcal{S}w(x) - \{\mathcal{S}\eta(x)\} = (s * u * v)(x) = (\mathcal{S}u * v)(x) = (u * \mathcal{S}v)(x), \qquad (2)$$

согласно которому с помощью ЛС-преобразований можно решать следующие задачи:

1. Используя отклик $w(x)$ ЛС-системы (с неизвестной передаточной функцией) на входной сигнал $v(x)$ реконструировать ее отклик $\mathcal{S}w(x)$ на входной сигнал $\mathcal{S}v(x)$.
2. Используя отклик $w(x)$ ЛС-системы с передаточной функцией $u(x)$ смоделировать отклик $\mathcal{S}w(x)$ ЛС-системы с передаточной функцией $\mathcal{S}u(x)$ на такой же входной сигнал.
3. Используя отклик $w(x)$ ЛС-системы с передаточной функцией $u(x)$ на входной сигнал $v(x)$ смоделировать отклик $\mathcal{S}_2\mathcal{S}_1w(x)$ ЛС-системы с передаточной функцией $\mathcal{S}_1u(x)$ на входной сигнал $\mathcal{S}_2v(x)$.

При использовании ЛС-преобразований никаких ограничений на вид или тип сигналов и функций кроме их интегрируемости по Лебегу на рассматриваемой области D не накладывается, что позволяет обрабатывать практически любые сигналы. Согласно ЛСП-уравнению (2) погрешность при ЛС-преобразованиях реальных сигналов составляет $\mathcal{S}\eta(x)$.

Представляя многомерные функции $f(x)$ в дискретном виде: $f = \sum_{i_1i_2\ldots} f_{i_1i_2\ldots}\delta_{i_1i_2\ldots}$, где $\delta_{i_1i_2\ldots}$ – символ Кронекера, отличающийся от 0 и равный 1 только на i_1i_2 ...-ой части разбиения области D, для ЛС-преобразований $g(x) = \mathcal{S}f(x) = (s * f)(x)$ в дискретном виде будем иметь:

$$g = \sum_{m_1m_2\ldots} g_{m_1m_2\ldots}\delta_{m_1m_2\ldots} = \sum_{i_1i_2\ldots}\sum_{j_1j_2\ldots} s_{i_1i_2\ldots}f_{j_1j_2\ldots}\delta_{i_1+j_1,i_2+j_2,\ldots}\,. \qquad (3)$$

Подставляя в (3) $i_k = m_k - j_k$ или $j_k = m_k - i_k$, для элемента $g_{m_1m_2\ldots}$ получим: $g_{m_1m_2\ldots} = \sum_{j_1j_2\ldots} s_{m_1-j_1,m_2-j_2,\ldots}f_{j_1j_2\ldots} = \sum_{i_1i_2\ldots} s_{i_1i_2\ldots}f_{m_1-i_1,m_2-i_2,\ldots}$. В одномерном случае: $g = \sum_i\sum_j s_iv_j\delta_{i+j}$ и $g_m = \sum_j s_{m-j}f_j = \sum_i s_if_{m-i}$.

Спектральная форма ЛС-преобразований имеет простой вид: $G(\omega) = S(\omega)F(\omega)$, или для дискретных спектров: $G_{i_1i_2\ldots} = S_{i_1i_2\ldots}F_{i_1i_2\ldots}$ ($G_i = S_iF_i$ в одномерном случае), но для существования спектров функции должны быть квадратично интегрируемы в области D, и если область D ограничена, то по умолчанию предполагается, что функции периодически продолжены вовне области D, что при, как правило, неопределенных граничных условиях резко ограничивает область применения спектральных методов. С физической точки зрения, прямые ЛС-преобразования дают неустановившиеся решения, а спектральные – установившиеся.

2. Базовые задачи ЛС-преобразований

Методы ЛС-преобразований относятся к прямым методам реконструкции, но их можно успешно использовать при решении отдельных задач, характерных для важных и часто встречающихся на практике случаев, даже когда большинство существующих математических методов оказываются неэффективными. Рассмотрим решение таких задач, которые будем называть базовыми, на примере реконструкции одномерных сигналов ЛС-систем. Решения могут быть расширены на случай многомерных сигналов и изображений.

Задача 1 (Два импульса). Пусть известен отклик $w_0(x)$ некоторой ЛС-системы (1) от двух смещенных относительно друг друга на расстояние (или время) $T > 0$ импульсов $v(x)$: $v_0(x) = v(x) - av(x-T)$ ($w_0(x) = v_0(x) = 0$ при $x < 0$). На заданном интервале $x \in (-\infty, L)$, $L > T$, реконструировать отклик $w(x)$ этой системы от импульса $v(x)$.

ЛС-система удовлетворяет уравнению (1): $w_0(x) = (u * v_0)(x) = (w * s)(x)$, $s(x) = \delta(x) - a\delta(x-T)$. Применив преобразования $\mathcal{S} = \mathcal{S}_n \cdots \mathcal{S}_1$, где $\mathcal{S}_k$: $\delta(x) + a^{2^{k-1}} w(x - 2^{k-1}T)$, к функции $s(x)$, получим: $\mathcal{S}s(x) = \delta(x) - a^{2^n}\delta(x - 2^nT)$. Тогда, согласно (2): $\mathcal{S}w_0(x) = w(x) - a^{2^n} w(x - 2^nT)$. Если $2^nT > L$, то на интервале $x \in (-\infty, L)$: $\mathcal{S}w_0(x) = w(x)$ (см. рисунок 1) и задача решена.

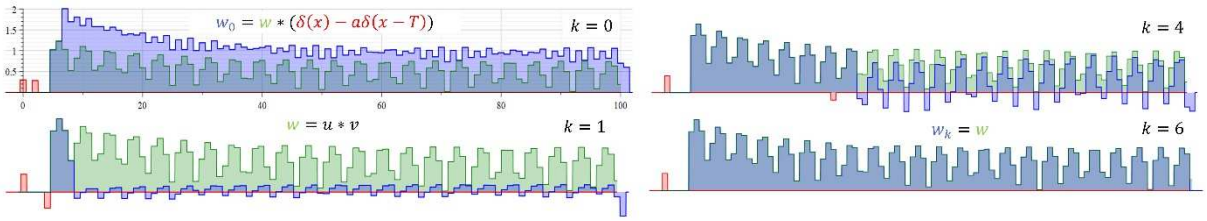


Рисунок 1 – Реконструкция сигналов при наложении двух импульсов

При наличии неопределенностей $\eta(x)$ погрешность $\mathcal{S}_k \cdots \mathcal{S}_1 \eta(x)$ на интервале $[2^{k-1}T, 2^kT)$ не превысит: $\eta(x)\left(1 + \sum_{i=1}^{k} a^{2^i}\right)$. При хорошей обусловленности задачи ($|a| < 1$) прирост погрешности a^{2^k} быстро уменьшается, а при плохой ($|a| > 1$) – возрастает (см. Таблицу 1).

Таблица 1 Прирост погрешности $\boldsymbol{a^{2^k}}$ на $\boldsymbol{k}$-том шаге итерации

$k \setminus a$	0.5	0.9	1.1	1.5
3	3.9e-3	4.3e-1	2.14	25.6
7	2.9e-39	1.4e-6	1.9e5	3.5e22

Таким образом за $n = [\log_2(L/T)] + 1$ итераций (или меньше, если прирост погрешности будет пренебрежимо мал) можно на произвольном интервале $x \in [0, L)$ реконструировать отклик любой ЛС-системы от единичного импульса при наложении двух смещенных входных импульсов. Погрешность уменьшается при уменьшении числа итераций, и, следовательно, при увеличении расстояния T между импульсами. #

Задача 2 (Прямоугольный импульс (изменение длительности)). Пусть известен отклик $w_0(x)$ произвольной ЛС-системы (1) на входной прямоугольный импульс длительности T: $v_0(x) = 1$ при $x \in [0, T]$ ($w_0(x) = v_0(x) = 0$ при $x < 0$). На заданном интервале $x \in [0, L)$ реконструировать отклик $w(x)$ этой системы от прямоугольного импульса длительности $t < T$: $v(x) = 1$ при $x \in [0, t]$.

Применяя ЛС-преобразование $\mathcal{S}_0$: $\delta(x) - \delta(x-t)$ к уравнению $w_0(x) = (u * v_0)(x)$, получим: $\mathcal{S}_0 w_0(x) = (u * \mathcal{S}_0 v_0)(x) = \Big(u * \big(v(x) - v(x-T)\big)\Big)(x)$, и перейдем к решению задачи 1 о наложении двух смещенных импульсов с неопределенностью $\eta(x)$ при $0 \le x < t$ и $\eta(x) - \eta(x-t)$ при $x \ge t$. Погрешность реконструкции будет минимальна при совпадении нулей спектров $V_0(\omega) = 2\sin(\omega T/2)/(\omega T/2)$ и $V(\omega) = 2\sin(\omega t/2)/(\omega t/2)$ сигналов $v_0(x)$ и $v(x)$, т.е. при $T = mt$, $m \in \mathbb{N}$. Когда $T > L$ функция $v_0(x)$ является ступенчатой на интервале $x \in (-\infty, L)$ и для реконструкции понадобится лишь одно ЛС-преобразование $\mathcal{S}_0$.

Задача 3 (Трапецеидальный импульс). Пусть известен отклик $w_0(x)$ произвольной ЛС-системы (1) на входной импульс: $v_0(x) = x/t$ при $x \in [0, t)$, $v_0(x) = 1$ при $x \in [t, T]$ и $v_0(x) = 1 - (x-T)/t$ при $x \in [T, T+t]$ ($w_0(x) = v_0(x) = 0$ при $x < 0$). На заданном интервале $x \in (-\infty, L)$ реконструировать отклик $w(x)$ этой системы от прямоугольного импульса длительности t: $v(x) = 1$ при $x \in [0, t]$.

Применяя ЛС-преобразование дифференцирования $\mathcal{S}_0$: d/dx к уравнению $w_0(x) = (u * v_0)(x)$, получим: $\mathcal{S}_0 w_0(x) = (u * \mathcal{S}_0 v_0)(x) = \Big(u * \big(v(x) - v(x-T)\big)\Big)(x)$, и перейдем к решению задачи 1, неопределенность которой будет: $\frac{d}{dx}\eta(x)$. Согласно Фурье-преобразованию спектр этой неопределенности возрастает как $FT\left[\frac{d}{dx}\eta(x)\right] = -I\omega FT[\eta(x)]$, где I – мнимая

единица. Погрешность уменьшается с увеличением длительности импульса T и фронта t и будет минимальна при $t > L$, что соответствует линейно возрастающей на интервале $x \in [0, L)$ функции $v(x)$. #

Задача 4 (Экспоненциальный импульс). Пусть известен отклик $w_0(x)$ произвольной ЛС-системы (1) на входной экспоненциальный импульс: $v_0(x) = \exp(-x/T)$, $T > 0$, при $x \geq 0$, ($w_0(x) = v_0(x) = 0$ при $x < 0$). На заданном интервале $x \in (-\infty, L)$ реконструировать отклик $w(x)$ этой системы от импульса $v(x) = v_0(x)$ при $x \in [0, t]$.

Применяя ЛС-преобразование $\mathcal{S}_0$: $\delta(x) - \exp(-t/T)\delta(x - t)$ к уравнению $w_0(x) = (u * v_0)(x)$, получим: $\mathcal{S}_0 w_0(x) = (u * \mathcal{S}_0 v_0)(x) = (u * v)(x) = w(x)$. Погрешность решения равна $\eta(x)$ при $x < t$ и $\eta(x) - \exp(-t/T)\eta(x - t)$ при $x \geq t$.

Задача 5. (Дискретные наложения импульсов) Пусть известен отклик $w_0(x)$ ЛС-системы на входной сигнал $v_0(x) = (v * s)(x)$, $s(x) = \sum_{k=0} s_k \delta(x - kt)$, $s_0 \neq 0$, $t > 0$, $w_0(x) = v_0(x) = 0$ при $x < 0$. На интервале $x \in (-\infty, L)$ реконструировать отклик $w(x)$ от импульса $v(x)$.

Применяя к функции $s(x)$ ЛС-преобразования $\mathcal{S} = \mathcal{S}_n \cdots \mathcal{S}_1$, где $\mathcal{S}_k$: $\delta(x) - (s_{k+1}^k / s_0)\delta(x - kt)$ и $s^k(x) = \mathcal{S}_k \cdots \mathcal{S}_1 s(x) = \sum_i s_i^k \delta(x - it)$, последовательно обнуляем коэффициенты s_{k+1}^k в функциях $s^k(x)$. На n-том шаге итерации при $nt > L$ на интервале $x \in (-\infty, L)$ получим: $s^n(x) = \mathcal{S}s(x) = s_0\delta(x)$. Отсюда, согласно (2): $\mathcal{S}w_0(x) = s_0 w(x)$. Погрешность решения на интервале $x \in [kt, kt + t)$ составит: $\mathcal{S}_k \cdots \mathcal{S}_1 \eta(x)$.

Отметим, что для дискретных сигналов с шагом дискретизации t решение задачи 5: $w_{i+1} = \left(w_i - \sum_{j=1}^{i} s_{i-j} w_j\right)/s_0$, $i = 0..n$, соответствует решению линейной системы уравнений с матрицей Теплица обратным ходом метода Гаусса-Жордана.

3. Примеры использования ЛС-преобразований

На Рис.2 представлен пример применения метода ЛС-преобразований при коррекции данных нейтронного спектрометра по времени пролета в Институте ядерных исследований РАН. Пучок нейтронов инициируется в нейтронной мишени коротким импульсом протонов из линейного ускорителя. Нейтроны разделяются по энергиям на пролетной базе и взаимодействуют с ядрами исследуемого вещества с испусканием γ–квантов, в результате чего γ–детектором регистрируется энергетический спектр ядер. Разрешающая способность спектрометра напрямую зависит от длительности импульса протонов и от длины пролетной базы. Но сформировать предельно короткие импульсы протонов непросто из-за проблем со стабилизацией пучка, что приводило к различным искажениям импульсов, например, к появлению после-импульсов, которые отчетливо видны во второй серии измерений, представленных на рисунке 2. С помощью метода ЛС-преобразований (базовая задача 1) искажения во второй серии измерений от после-импульса были успешно устранены всего за три итерации и были получены реальные спектры исследуемого вещества (подробнее см. [8]).

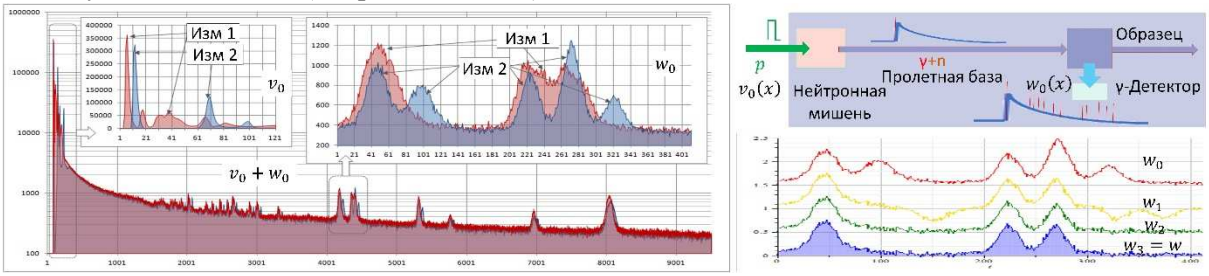


Рисунок 2 – Корректировка данных нейтронного спектрометра ИЯИ РАН

Метод ЛС-преобразований позволяет повысить эффективную разрешающую способность спектрометра. Пусть для достижения необходимой разрешающей способности спектрометра требуется импульс пучка протонов длительностью t, но стабильная проводка пучка обеспечивается лишь при длительности T большей t. Проводя измерения при длительности T с дальнейшей реконструкцией данных спектрометра от импульса меньшей длительности t с помощью решения базовой задачи 2 можно обеспечить и стабильную работу, и необходимую разрешающую способность спектрометра без проведения трудоемких и дорогостоящих работ по модернизации оборудования и конструкции спектрометра. При этом фронты пучка должны быть

меньше t, т.к. в противном случае упадет амплитуда импульса. Это ограничение можно устранить с помощью решения базовой задачи 3, но необходимо учитывать повышение уровня шумов при такой реконструкции.

С помощью метода ЛС-преобразований можно осуществлять реконструкцию сигналов любых ЛС-систем. На рисунке 3 представлен пример реконструкции отклика емкостного делителя напряжения на трапецеидальный сигнал длительностью 30 нс из его отклика на трапецеидальный сигнал большей длительности 120 нс с помощью решения базовой задачи 2. Представленные на графиках исходные сигналы были рассчитаны с помощью симулятора MicroCAP от Spectrum Software. Для реконструкции сигналов на интервале от 0 до 500 нс потребовалось всего три итерации. Погрешность реконструкции 0.0012 % обусловлена в основном погрешностью дискретизации сигналов при численных расчетах, а не погрешностью метода ЛС-преобразований.

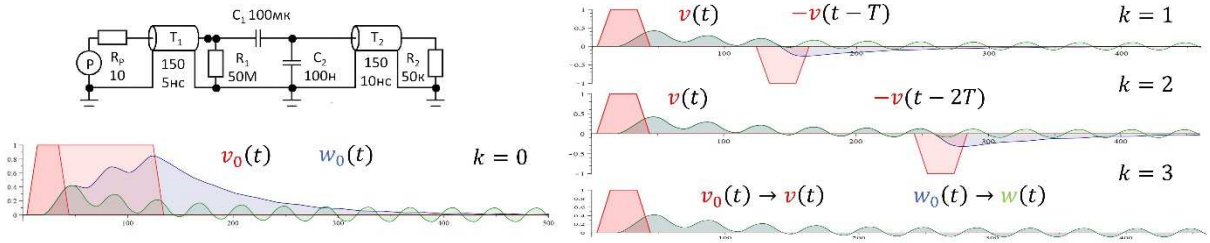


Рисунок 3 – Моделирование отклика ЛС-системы

С помощью метода ЛС-преобразований можно осуществлять реконструкцию не только одномерных, но и многомерных сигналов. На рисунке 4 представлен процесс реконструкции двумерного изображения с помощью решения базовой задачи 5 при произвольном размытии изображения. Для простоты рассмотрен случай размытия вдоль горизонтальной оси, но реконструкция возможна и при размытии в результате произвольных плоско-параллельных смещений изображения.

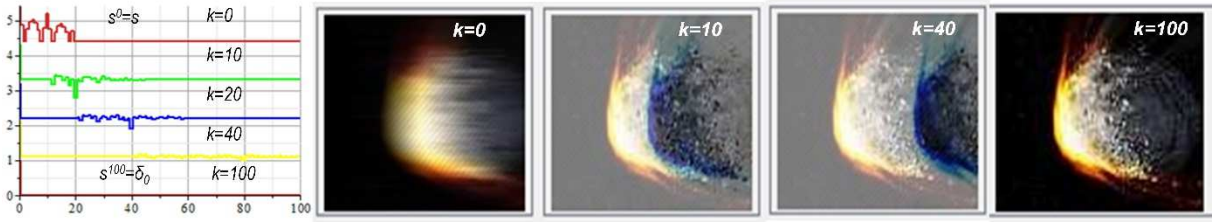


Рисунок 4 – Реконструкция изображений

Методы ЛС-преобразований можно использовать для достижения сверх разрешения сигналов и изображений. Последовательно сгруппированные сигналы n приборов с разрешением T, смещенных относительно друг друга на шаг $t = T/n$, составляют ЛС-уравнение (1). Наиболее эффективно при этом использовать искажения сигналов прямоугольного или экспоненциального вида из базовых задач 2 или 4, т.к. реконструкцию можно осуществить за один шаг итерации. В дискретном виде решение базовой задачи 4 можно использовать для регистрации и разделения сигналов в детекторах частиц большой интенсивности (см., например, [9,10]).

Заключение

Предлагаемые в данной работе математические методы ЛС-преобразований предназначены для коррекции, реконструкции и моделирования сигналов и изображений, искаженных в результате их обработки линейными стационарными системами. ЛСП-методы позволяют осуществить реконструкцию и моделирование сигналов при минимуме исходной информации о сигналах и системах обработки, применимы для обработки как дискретных, так и непрерывных сигналов, а также многомерных сигналов. Методы относятся к методам прямой инверсии и практически не ограничивают класс обрабатываемых функций, что позволяет обрабатывать кусочно-непрерывные функции с разрывами первого рода, сигналы спектрометров с узкими спектральными линиями, контрастные изображения с резкими границами и т.п.

Прямые методы не подходят для решения некорректно поставленных и плохо обусловленных задач реконструкции сигналов, а приближенные методы, такие как техника регуляризации, неэффективны при восстановлении функций такого широкого класса, поскольку приводят к значительным искажениям сигналов. ЛСП-методы в отличие от других прямых методов в ряде важных и часто встречающихся на практике случаях позволяют найти решение с минимальными погрешностями за счет выбора типа ЛС-преобразований и сокращения их числа.

ЛСП-методы позволяют значительно улучшить точность систем обработки и их разрешающую способность вплоть до сверх разрешения. Предлагаемые методы являются доступной альтернативой аппаратным методам, связанным с решением сложных научных и технических проблем и с созданием дорогостоящего оборудования. Методы ЛС-преобразований имеют широкие перспективы практического применения для повышения эффективности работы как различных экспериментальных приборов, так и целых установок.

Литература

1. Mueller J.L., Siltanen S. Linear and Nonlinear Inverse Problems with Practical Applications. SIAM, vol.10, 2012.
2. Hansen P.C. Discrete Inverse Problems: Insight and Algorithms. Fundamentals of Algorithms. Philadelphia: SIAM, 2010.
3. Vogel C.R. Computational methods for inverse problems. Philadelphia: SIAM, 2002.
4. Hespanha J.P. Linear System Theory. Princeton university press, 2009.
5. Vladimirov V.S. Equations of Mathematical Physics. Moscow: Nauka, 1981.
6. Kazufumi I., Bangti J. Inverse Problems: Tikhonov Theory and Algorithms. Singapore: World Scientific, 2014.
7. Elsayed M.A., El-Shafai W., Rashwan M.A., Dessouky M.I., El-Fishawy A.S., Abd El-Samie F.E. Efficient iterative implementation of regularized solutions for image and signal reconstruction problems. Journal of Optics, 2023. *https://doi.org/10.1007/s12596-023-01179-w*.
8. Новиков-Бородин А.В. Реконструкция экспериментальных данных спектрометров по времени пролёта методом пошагового сдвига. Приборы и техника эксперимента, 2018, № 6, с. 17–24. *https://doi.org/10.1134/S0032816218050269*.
9. V.T. Jordanov, Unfolding-synthesis technique for digital pulse processing, part 1: Unfolding. Nucl. Instrum. Methods Phys. Res. A 805, 63-71, 2016. doi: *https://doi.org/10.1016/j.nima.2015.07.040*
10. V.T. Jordanov, K.V. Jordanova, Unfolding-synthesis technique for digital pulse processing, Part 2: Synthesis. Nucl. Instrum. Methods Phys. Res. A 1004, 167421, 2022. *https://doi.org/10.1016/j.nima.2022.167421*.

THE USE OF LS-TRANSFORMATIONS FOR RECONSTRUCTION, CORRECTION AND MODELING OF EXPERIMENTAL DATA

A.V. Novikov-Borodin
Institute for Nuclear Research of RAS, 117312 Moscow, Russia

Abstract. The possibilities of using the mathematical methods of LS-transformations for processing of experimental data distorted by linear stationary systems are considered. The methods allow correction, reconstruction and modeling of processing system signals in important and frequently encountered in practice cases when most of the existing both direct and approximate mathematical methods are ineffective. Examples of improving the effective quality parameters of experimental equipment, including electronic instruments, receiving and transmitting devices, spectrometers, up to ultra-high resolution are considered. The proposed methods make it possible to improve the quality of data processing without using complex and expensive hardware methods of equipment modification.

Keywords: mathematical processing of signals and images; LS-transformations technique; linear stationary systems; multidimensional convolution type equation.

АНАЛИЗ ЭФФЕКТИВНОСТИ ПРИМЕНЕНИЯ НЕЙРОННОЙ СЕТИ ДЛЯ ЗАДАЧИ ЦИФРОВОГО ПРЕДЫСКАЖЕНИЯ СИГНАЛА

Аспирант Хоанг Ван Туан

Московский физико-технический институт

В данной работе рассматривается задача цифрового предыскажения для компенсации нелинейных искажений и эффектов памяти усилителя мощности. Сравниваются два подхода: традиционная полиномиальная модель памяти и нейронная сеть прямого распространения. Для анализа были выполнены эксперименты в среде MATLAB, эффективность предыскажения оценивалась по двум показателям: нормализованной среднеквадратичной ошибке NMSE и коэффициенту мощности соседнего канала ACPR. Результаты моделирования показали, что оба метода DPD успешно улучшают линейность усилителя мощности и компенсируют эффекты памяти. Однако нейронная сеть продемонстрировала более высокую эффективность по сравнению с полиномиальной моделью, обеспечив лучшее подавление мощности в соседних каналах и более низкое значение NMSE.

Введение

В современных беспроводных системах связи усилитель мощности (УМ) играет ключевую роль в усилении радиочастотного сигнала перед его передачей на антенну. Однако УМ часто сталкивается с проблемами нелинейности, особенно при работе вблизи зоны насыщения, что приводит к значительным искажениям сигнала. Эти нелинейные искажения не только снижают качество сигнала в передающем канале, но и создают помехи в соседних каналах, что отрицательно сказывается на эффективности использования частотного спектра и производительности всей системы. Помимо нелинейности, УМ подвержен эффектам памяти, при которых выходные характеристики зависят не только от текущего входного сигнала, но и от предшествующих значений. Эффекты памяти могут быть вызваны физическими факторами, такими как тепловые характеристики, накопление заряда или элементы схемы согласования импедансов в конструкции УМ. Эти эффекты значительно усложняют процесс моделирования и линеаризации УМ, требуя применения более совершенных методов обработки сигналов.

Метод цифрового предыскажения (Digital Predistortion – DPD) [1,2] зарекомендовал себя как высокоэффективное решение, обеспечивающее оптимальное соотношение затрат и производительности. DPD предполагает использование нелинейного предыскажателя, функционально инверсного к характеристикам УМ, что позволяет компенсировать проявления нелинейных эффектов и эффектов памяти. Это, в свою очередь, способствует значительному повышению линейности УМ и улучшению качества сигнала в системах связи.

Разработка эффективного DPD требует обширных знаний нелинейности и эффекта памяти блока УМ [3]. Существуют две популярные архитектуры для проектирования DPD: Архитектура косвенного обучения (Indirect Learning Architecture – ILA) [1] и архитектура прямого обучения (Direct Learning Architecture – DLA) [2]. Архитектура DLA предполагает прямую оптимизацию параметров предыскажения с учётом выходного сигнала усилителя мощности, что позволяет эффективно минимизировать искажения. Однако этот подход имеет несколько значительных недостатков, среди которых высокая вычислительная сложность. Адаптивные алгоритмы, используемые в DLA, требуют множества итераций для нахождения набора параметров, который минимизирует критерий оптимизации [2]. Это значительно увеличивает время вычислений и усложняет процесс идентификации. В связи с этими ограничениями архитектура ILA становится более предпочтительным выбором для решения задач цифрового предыскажения. ILA позволяет более эффективно идентифицировать параметры блока предыскажителя с меньшей вычислительной сложностью, что делает её более применимой в условиях ограниченных вычислительных ресурсов.

С развитием нейронных сетей для задач поведенческого моделирования УМ и DPD стали активно использоваться модели на основе рекуррентных (RNN) и сверточных нейронных сетей (CNN) [4,5,6,7]. Однако такие методы имеют значительные недостатки: высокая вычислительная сложность и требовательность к ресурсам, что ограничивает их применение в задачах реального времени. На этом фоне сети прямого распространения (FNN) представляют собой перспективную альтернативу, способную обеспечить сравнимую точность линеаризации при значительно меньших вычислительных затратах.

Данное исследование направлено на анализ и сравнение эффективности поведенческого моделирования УМ и DPD при применении архитектуры косвенного обучения с алгоритмом адаптации RLS (рекурсивный алгоритм наименьших квадратов) [8] и FNN.

Модель цифрового предыскажения сигнала на основе архитектуры косвенного обучения

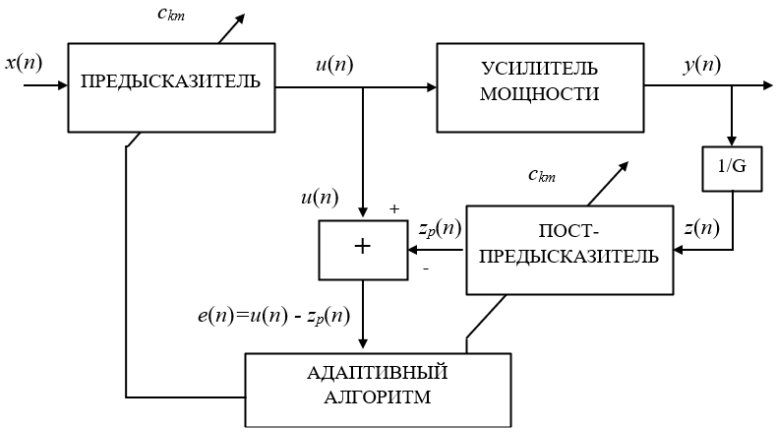


Рисунок 1. Архитектура косвенного обучения (ILA)

На рисунке 1 представлена схема DPD на основе архитектуры ILA с адаптивным алгоритмом RLS. Сигнал $x(n)$ является входным сигналом предыскажителя; $u(n)$, $y(n)$ являются входным и выходным сигналами УМ; $z(n) = y(n)/G$; G – коэффицент усиления; $z_p(n)$ – выходной сигнал блока пост-предыскажителя; сигнал ошибки $e(n) = u(n) - z_p(n)$; c_{km} – коэффициенты предыскажителя. Взаимосвязь между входным и выходным сигналом модели предыскажителя математически описывается следующим образом:

$$u(n) = \sum_{k=1}^{K}\sum_{m=0}^{M} c_{km} x(n-m)\,|\,x(n-m)\,|^{k-1}, \tag{1}$$

где $c_{km} = [c_{10},...,c_{K0},...,c_{1M},...,c_{KM}]^T$; K, M – порядок нелинейности и глубина памяти.

Вход и выход модели пост-предыскажителя УМ можно выразить следующими уравнениями:

$$z_p(n) = \sum_{k=1}^{K}\sum_{m=0}^{M} c_{km} z(n-m)\,|\,z(n-m)\,|^{k-1}, \tag{2}$$

$$y(n) = \sum_{k=1}^{K}\sum_{m=0}^{M} a_{km} u(n-m)\,|\,u(n-m)\,|^{k-1}, \tag{3}$$

где a_{km} – коэффициенты МР модели усилителя мощности.

Алгоритм RLS, использующий экспоненциальный фактор забывания $\lambda \le 1$, минимизирует целевую функцию:

$$\xi(n) = \sum_{i=0}^{n} \lambda^{n-i} e^2(i). \tag{4}$$

Общий алгоритм RLS описывается следующим образом:

1. Определить мгновенную ошибку $e(n) = u(n) - z_p(n)$.

2. Обновить матрицу весов

$$k(n) = \frac{P(n-1)z(n)}{\lambda + z^T(n)P(n-1)z(n)}, \tag{5}$$

$$P(n) = \frac{1}{\lambda}\Big[P(n-1) - k(n)z^T(n)P(n-1)\Big]. \tag{6}$$

3. Обновить коэффициенты предыскажителя

$$c(n) = c(n-1) + k(n)e(n). \tag{7}$$

Модель цифрового предыскажения сигнала на основе нейронной сети

Особенность нейронных сетей заключается в том, что они обучаются на основе данных входных и выходных сигналов УМ, автоматически подстраивая свои параметры под характерные особенности нелинейности и памяти системы. При этом не требуется заранее определять структуру модели, что делает нейронные сети гибким инструментом для широкого спектра задач.

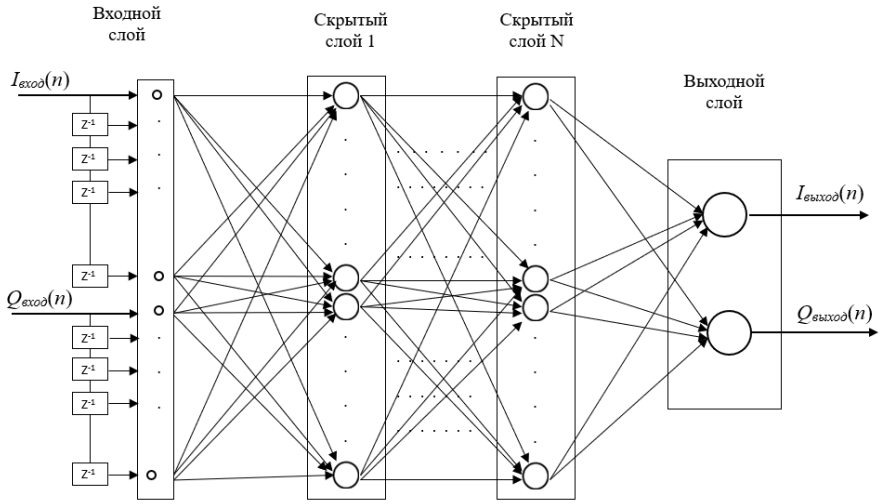


Рисунок 2. Архитектура обучения нейронной сети FNN

Типичная архитектура FNN (Рисунок 2) включает входной слой, принимающий сигнал УМ, скрытые слои с нелинейными функциями активации и выходной слой, прогнозирующий сигнал после предыскажения.

На входной слой подаются два канала сигнала: действительная часть $I_{вход}(n)$ и мнимая часть $Q_{\text{вход}}(n)$. Чтобы учесть эффекты памяти УМ, входные сигналы дополнительно проходят через цепочку задержек Z^{-1}.

$$X_{вход}(n) = \begin{bmatrix} I_{вход}(n), & I_{вход}(n-1), & \ldots, & I_{вход}(n-M), \\ Q_{вход}(n), & Q_{вход}(n-1), & \ldots, & Q_{вход}(n-M) \end{bmatrix}^T, \quad (8)$$

где M – глубина памяти (количество задержек).

Каждый нейрон слоя k вычисляет линейную комбинацию входных данных:

$$Z_j^{(k)}(n) = \sum_{i=1}^{N_{k-1}} W_{ji}^{(k)} A_i^{(k-1)}(n) + b_j^{(k)}, \quad (9)$$

где $Z_j^{(k)}(n)$ – входное значение нейрона j слоя k; $A_i^{(k-1)}(n)$ – выход i-го нейрона предыдущего слоя (или элемент $X_{вход}(n)$ если $k = 1$); $W_{ji}^{(k)}$ – весовая связь между i-м нейроном слоя $k-1$ и j-м нейроном слоя k; $b_j^{(k)}$ – смещение (bias) нейрона j слоя k; N_{k-1} – количество нейронов в предыдущем слое.

Нейрон применяет нелинейную функцию активации к своему входному значению:

$$A_j^{(k)}(n) = f\left(Z_j^{(k)}(n)\right) \quad (10)$$

Выход каждого скрытого слоя представлен вектором:

$$A^{(k)}(n) = \left[A_1^{(k)}(n), \quad A_2^{(k)}(n), \quad \ldots, \quad A_{N_k}^{(k)}(n) \right]^T, \quad (11)$$

где N_k – количество нейронов в слое k.

Выходной слой преобразует результат последнего скрытого слоя $A^{(N)}(n)$ в прогнозируемые значения $I_{\text{выход}}(n)$ и $Q_{\text{выход}}(n)$. Каждый выходной нейрон вычисляет:

$$Z_j^{\text{выход}}(n) = \sum_{i=1}^{N} W_{ji}^{\text{выход}} A_i^{(N)}(n) + b_j^{\text{выход}}, \quad (12)$$

где $Z_j^{\text{выход}}(n)$ – входное значение j-го нейрона выходного слоя; $A_i^{(N)}$ – выход i-го нейрона последнего скрытого слоя; $W_{ji}^{\text{выход}}$ – весовая связь между i-м нейроном последнего скрытого слоя и j-м выходным нейроном; $b_j^{\text{выход}}$ – смещение выходного нейрона. Выходной нейрон применяет функцию активации $Y_j(n) = f\left(Z_j^{\text{выход}}(n)\right)$.

Выходной вектор:

$$Y(n) = \left[I_{\text{выход}}(n), \quad Q_{\text{выход}}(n) \right]^T. \quad (13)$$

Результат моделирования

Для оценки эффективности поведенческой модели и применения DPD используются два показателя: NMSE (нормированная среднеквадратичная ошибка) и ACPR (коэффициент мощности соседнего канала).

NMSE является критерием оценки, показывающим общее отклонение между предсказанными и измеренными значениями во временной области. Формула NMSE представлена следующим образом:

$$NMSE[dB] = 10\log_{10}\left(\frac{\sum_{i=1}^{N} |y_i - \hat{y}_i|^2}{\sum_{i=1}^{N} |y_i|^2} \right) \text{dB}, \quad (14)$$

где y_i – реальное значение, а $\hat{y}_i$ – прогнозируемое значение.

ACPR – это критерий оценки в частотной области, который измеряет соотношение мощности в соседних каналах к мощности в основном канале.

Формула для вычисления ACPR:

$$ACPR = 10\log_{10}\left(\frac{P_{\text{adj}}}{P_{\text{main}}} \right) \text{dB}, \quad (18)$$

где P_{adj} – мощность в соседнем канале, P_{main} – мощность в основном канале.

С использованием программы MATLAB для моделирования модели MP с параметрами: $K = 7$, $M = 3$, $\lambda = 0{,}999$ и начальным вектором коэффициентов $c(0) = [1, 0, \ldots, 0]^T$. Для нейронной сети FN структура сети включает 1 входной слой, 2 скрытых слоя и 1 выходной слой.

На рисунках 3 и 4 представлены характеристики AM/AM и AM/PM УМ при использовании цифрового предыскажения на основе MP и FNN. Обе модели эффективно решили задачу линеаризации усилителя мощности и компенсации эффектов памяти.

Рисунок 5 показывает спектральную плотность мощности (PSD) на выходе усилителя мощности. Видно, что оба метода DPD, основанные на MP и FNN, значительно улучшают спектральную плотность мощности выходного сигнала УМ. Однако DPD с использованием FNN обеспечивает лучшую производительность, снижая мощность утечек в соседние каналы сильнее по сравнению с методом MP.

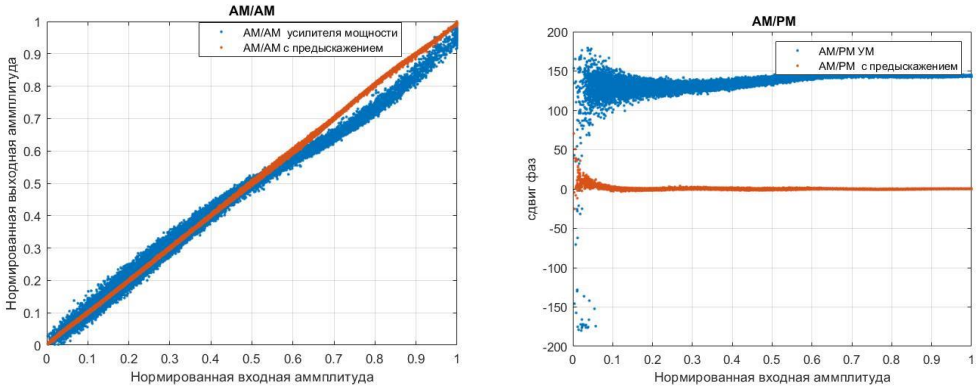


Рисунок 3. AM/AM, AM/PM характеристика при применении MP

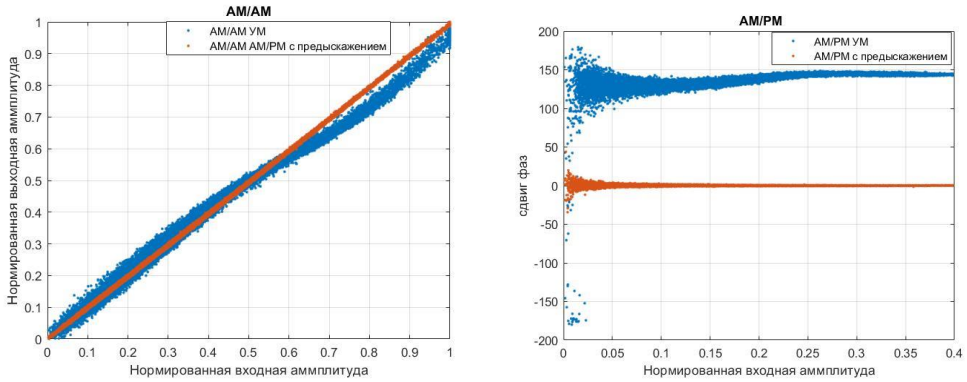


Рисунок 4. AM/AM, AM/PM характеристика при применении FNN

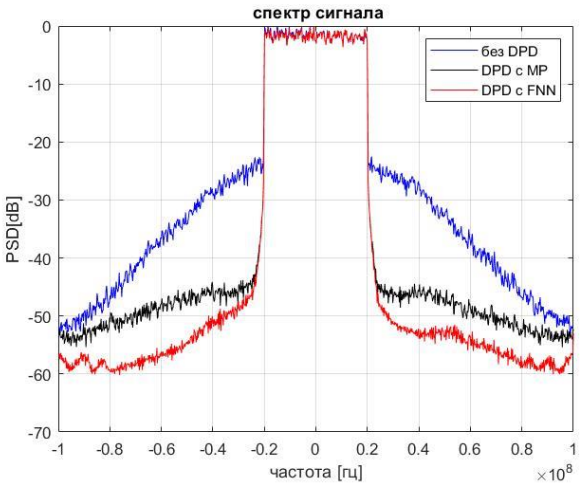


Рисунок 5. Спектральная мощность (PSD) на выходе УМ без и с применением DPD

Таблица 1. Сравнение NMSE [дБ], ACPR [дБ] при применении методов MP и FNN

Метод	**NMSE [дБ]**	**ACPR [дБ]**
MP	-35,48	-43,59
FNN	-39,69	-51,29

Заключение

В данной работе был проведен анализ и сравнение двух подходов к цифровому предыскажению сигнала для линеаризации усилителя мощности: MP и FNN. Моделирование в среде MATLAB показало, что оба метода эффективно решают задачу компенсации нелинейных искажений и эффектов памяти, характерных для усилителя мощности. Результаты сравнения на основе критериев NMSE и ACPR подтвердили, что метод FNN демонстрирует лучшие показатели по сравнению с MP. В частности, FNN обеспечивает более низкий уровень нормализованной ошибки и уменьшение мощности утечек в соседние каналы, что подтверждается значительным улучшением ACPR. Таким образом, нейронные сети представляют собой перспективное решение для задачи линеаризации УМ, особенно в условиях работы с широкополосными сигналами.

Литература

1. Borel A., Barzdėnas V., Vasjanov A. Linearization as a solution for power amplifier imperfections: A review of methods //Electronics. – 2021. – Т. 10. – №. 9. – С. 1073.
2. Singh S., Malik J. Review of efficiency enhancement techniques and linearization techniques for power amplifier //International Journal of Circuit Theory and Applications. – 2021. – Т. 49. – №. 3. – С. 762-777.
3. Lv J. et al. Full-Wave Digital Predistortion Linearization Based on Memory Polynomials for HF Power Amplifiers //IEEE Access. – 2024.
4. P. Jaraut et al., "Behavior Modeling and Digital Predistortion of Mismatched Wireless Transmitters Using Convolutional Neural Networks," in IEEE Transactions on Circuits and Systems II: Express Briefs, vol. 70, no. 1, pp. 336-340, Jan. 2023.
5. Javid-Hosseini S. H. et al. A Unified Neural Network-Based Approach to Nonlinear Modeling and Digital Predistortion of RF Power Amplifier //IEEE Transactions on Microwave Theory and Techniques. – 2024.
6. Wang W. et al. LSTM-CNN for Behavioral Modeling and Predistortion of 5G Power Amplifiers //2022 IEEE 9th International Symposium on Microwave, Antenna, Propagation and EMC Technologies for Wireless Communications (MAPE). – IEEE, 2022. – С. 28-32.
7. He, Ziming, and Fei Tong. "Residual RNN models with pruning for digital predistortion of RF power amplifiers." *IEEE Transactions on Vehicular Technology* 71.9 (2022): 9735-9750.
8. Le D. H. et al. Linearization of RF power amplifiers in wideband communication systems by adaptive indirect learning using RPEM algorithm //Mobile Networks and Applications. – 2020. – Т. 25. – С. 1988-1997.

ANALYSIS OF THE EFFICIENCY OF USING A NEURAL NETWORK FOR THE DIGITAL PREDISTORTION

Hoang Van Tuan

Moscow Institute of Physics and Technology

In this paper, the problem of digital predistortion for compensating nonlinear distortions and memory effects of a power amplifier is considered. Two approaches are compared: a traditional polynomial memory model and a feedforward neural network. For the analysis, experiments were performed in the MATLAB environment, the predistortion efficiency was estimated by two indicators: the normalized mean square error NMSE and the adjacent channel power ratio ACPR. The simulation results showed that both DPD methods successfully improve the linearity of the power amplifier and compensate for memory effects. However, the neural network demonstrated higher efficiency compared to MP, providing better power suppression in adjacent channels and a lower NMSE value.

ИССЛЕДОВАНИЕ МОДЕЛИ МАРШРУТИЗАЦИИ КАНАЛОВ ИНФОРМАЦИОННОГО ВЗАИМОДЕЙСТВИЯ В СЕТИ FANET

проф. Приоров А.Л., асп. Идиатуллин А.В., студ. Мороз О.Д.

Ярославский государственный университет им. П.Г. Демидова, г. Ярославль, Россия

С помощью комплекса программного обеспечения NS-3 проведено исследование и сравнение протоколов маршрутизации в беспроводной самоорганизующейся (БСОС) сети FANET. Для реализации задачи при моделировании использовались протоколы OLSR, AODV и DSDV. Определены параметры моделирования (число узлов, скорости, размеры области, тип трафика), а также описаны метрики, по которым оценивается производительность (Delay, Throughput, Overhead). Уточнен алгоритм проведения эксперимента, включая сбор логов, анализ pcap и использование FlowMonitor для автоматической статистики.

Введение

В настоящее время широкое распространение получил класс компьютерных сетей (беспроводных, децентрализованных), для которых характерна мобильность. Узлы таких сетей могут перемещаться в пространстве, включаться в сеть или исключаться из нее. Кроме того, для таких сетей характерно ограничение расстояния между узлами. Каждый узел характеризуется мощностью передатчика и чувствительностью приемника сигналов.

Алгоритм маршрутизации в таких сетях должен динамически находить маршрут, складывающийся из частичных маршрутов для каждого сообщения. Для этого могут использоваться знания, накопленные о системе и ее окружении. Поведение мобильных узлов трудно прогнозировать. Поэтому необходимо использовать аппарат теории графов и теории вероятностей (случайные графы). Для анализа информационных потоков в компьютерных сетях кроме методов, представленных теорией графов, используют также аналитические методы теории массового обслуживания и методы имитационного моделирования.

Имитационные модели являются одним из важнейших инструментов для исследования алгоритмов маршрутизации по той причине, что мобильные сети и поведение объектов мобильных сетей являются достаточно сложными и не всегда дают возможность применить аналитические методы. Поскольку математической моделью является динамический граф, то инструментальные средства имитационного моделирования должны быть гибкими и эффективными. В качестве таких инструментальных средств выбрана система NetworkSimulator-3 (NS-3).

Постановка задачи

При работе с FANET особое внимание уделяют описанию динамической структуры сети: в какой степени движение беспилотных летательных аппаратов (БЛА, БПЛА) отражает реальную кинематику, что послужит первоосновой для построения траекторий, какие режимы или профили полета заложены в модель. Данные вопросы предопределяют корректность интерпретации симуляционных результатов [1].

Для сети FANET важно также учесть высоту полета, скорость перемещения в трехмерном пространстве и потенциальное воздействие внешних факторов, например, влияние ветра и турбулентности, если такой уровень детализации необходим по условиям постановки конкретной задачи.

В дополнение к базовым настройкам существует набор показателей, которые необходимо оценить. Среди них пропускная способность (Throughput), задержка передачи пакетов (Delay), отношение числа доставленных пакетов к числу отправленных (Packet Delivery Ratio), а также величина служебного трафика, генерируемого протоколом (Overhead) [2].

В рассматриваемой задаче моделирования FANET требуется определить число узлов, их начальное расположение, сценарии движения и параметры работы каждой группы узлов. Если в сети присутствуют стационарные пункты (например, наземная станция управления или ретранслятор), то их координаты и характеристики передачи остаются неизменными в ходе эксперимента, тогда как беспилотные узлы перемещаются по заданной схеме [3].

Значимый элемент описания экспериментальной среды – тип нагрузки (trafficpattern). В сети FANET ее узлы часто обмениваются короткими сообщениями телеметрии или управляющими командами. Маршрутизационный протокол может показывать разную эффективность в зависимости от размера пакетов и интенсивности потока, особенно если речь идет о высоких скоростях передачи и интенсивных коллизиях на канальном уровне [4].

Задача моделирования, таким образом, состоит в создании симуляционных сценариев, где можно систематически менять ключевые входные параметры и наблюдать реакцию протокола маршрутизации на изменения в динамике сети. Кроме того, анализ полученных данных позволит сформулировать рекомендации по выбору наиболее эффективного подхода к маршрутизации в реальных приложениях [5].

Выделим три выбранных протокола, наиболее популярные при работе в беспроводных самоорганизующихся сетях: DSDV, AODV и OLSR. Их функциональные принципы существенно различаются, что позволяет выявить сильные и слабые стороны каждой концепции при быстром изменении топологии.

Протокол DSDV – проактивный протокол маршрутизации на основе дистанционно-векторного алгоритма Bellman-Ford. При работе DSDV каждый узел хранит маршрут до всех прочих узлов сети и периодически передает обновленную информацию о своих таблицах маршрутов соседям. Каждая запись в таблице маршрутов содержит идентификатор назначения, расстояние (число переходов) и метку последовательности (sequencenumber), указывающую актуальность. При получении обновления от соседа узел сопоставляет метки последовательности и, если принятая информация более свежая, корректирует маршрут [6].

Протокол AODV относится к семейству реактивных протоколов маршрутизации, где маршруты формируются «по требованию». Когда узлу необходимо передать пакет в неизвестный ему пункт назначения, он рассылает широковещательный запрос (RREQ) по сети. Если узел-получатель или промежуточный узел, располагающий маршрутом до получателя, принимает этот запрос, он отправляет ответ (RREP) обратно к источнику [7].

Протокол OLSR – проактивный протокол, основанный на концепции периодического распространения информации о состоянии связей (link state). Однако, в отличие от традиционного линк-стейт-протокола, OLSR оптимизирует широковещательные рассылки с помощью механизма MultiPoint Relays (MPR). Суть метода MPR состоит в том, что для каждого узла выбирается ограниченный набор соседей (MPR), которые транслируют сообщения о состоянии сети дальше. Это сокращает объем служебного трафика, поскольку не все узлы пересылают информацию, а лишь узлы-ретрансляторы в рамках выбранного маршрута, что существенно снижает общий уровень накладных расходов (Overhead).

В сети FANET протокол OLSR может показывать хорошие результаты при достаточно плотном расположении узлов, поскольку тогда выбор оптимального набора MPR может эффективно распространять свежие данные о топологии по всей сети. Однако если сеть расформирована, и узлы сильно рассредоточены, могут возникнуть проблемы с корректной поддержкой связности, особенно при высоком уровне мобильности, когда набор MPR постоянно меняется. Тем не менее, при благоприятных условиях протокол OLSR часто дает достаточно малые задержки, так как узлы заранее формируют карту топологии и готовы отправлять пакеты практически мгновенно.

Проведение эксперимента и результаты

Представлена реализация протоколов маршрутизации DSDV, AODV и OLSR в симуляционной среде NS-3, описываются ключевые аспекты создания сценариев, организации исходного кода и настройка параметров среды моделирования. В процессе разработки имитационного эксперимента использовался пакет NS-3 версии 3.35 под управлением операционной системы Ubuntu 22.04 LTS.

В ходе работы сформирован ряд требований к моделированию высокодинамичных сетей БЛА. Ключевые моменты, которые влияют на реализацию протоколов в NS-3, включают:

1. Описание мобильности: необходимо использовать модель RandomWaypointMobilityModel с заданными скоростями (5, 10, 15, 20 м/с) и областью перемещения 1500 × 1500 м.

2. Сетевую топологию: количество узлов (10, 20, 30) варьируется в зависимости от сценария.

3. Протоколы маршрутизации: реализуются три основных подхода (DSDV, AODV, OLSR) в их стандартной конфигурации, при этом для каждого протокола задаются дополнительные параметры (частота Hello-сообщений, период обновления маршрутных таблиц, таймеры).

4. Тип трафика: нужно смоделировать UDP-потоки (OnOff с битрейтом 512 кбит/с) и TCP-потоки (BulkSend) для оценки пропускной способности (Throughput) и задержек в зависимости от транспортного протокола.

В данном исследовании сопоставляются результаты, полученные при использовании протоколов DSDV, AODV и OLSR, с учетом изменений таких параметров, как скорость, их количество и тип трафика. Для наглядности приводятся таблицы, отражающие три ключевые метрики: задержка (End-to-End Delay), пропускная способность (Throughput) и накладные расходы (RoutingOverhead). Кроме того, дается интерпретация наблюдаемых закономерностей и проводится оценка причин, по которым один протокол может работать лучше другого при определенной конфигурации сети.

Первой метрикой является средняя задержка передачи данных (End-to-End Delay). Эта величина особенно важна для приложений реального времени – например, для передачи видеопотока или команд управления на БЛА. Ниже приводится примерная сводная таблица 1.

Таблица 1. Средняя задержка передачи данных (End-to-End Delay)

Число узлов БСОС	Скорость БЛА (м/с)	DSDV	AODV	OLSR
10	5	0,0489	0,0448	0,0482
10	10	0,0526	0,0544	0,0526
10	20	0,0653	0,0567	0,0491
20	5	0,0672	0,0767	0,0656
20	10	0,0971	0,0838	0,0671
20	20	0,1077	0,0753	0,0707
30	5	0,1307	0,0875	0,0859
30	10	0,1622	0,0976	0,0998
30	20	0,1797	0,1244	0,1501

Таблица показывает среднюю задержку передачи данных для протоколов DSDV, AODV и OLSR в зависимости от количества узлов и скорости движения в сети FANET. Увеличение скорости приводит к росту задержки во всех протоколах, особенно у DSDV, так как он использует периодическое обновление таблиц маршрутов. При низких скоростях (5 м/с) протоколы OLSR и AODV показывают лучшие результаты, тогда как протокол DSDV уступает им при увеличении количества узлов. При средних скоростях (10 м/с) протокол OLSR остается наиболее стабильным, протокол AODV демонстрирует умеренные задержки, а протокол DSDV значительно теряет эффективность. При высокой скорости (20 м/с) протокол OLSR показывает наименьшую задержку, тогда как протокол DSDV страдает от резкого роста задержек, особенно при 30 узлах. В целом, протокол OLSR оказывается наиболее эффективным для FANET в условиях высокой мобильности, протокол AODV обеспечивает баланс между задержкой и накладными расходами, а протокол DSDV подходит только для стабильных сетей с низкой подвижностью.

При небольшом числе узлов (10) и низкой скорости (5 м/с) задержки у протоколов DSDV, AODV и OLSR находятся примерно в одном диапазоне (0,04–0,05 с), причём протокол AODV иногда показывает самую низкую задержку, что может быть связано с небольшой необходимостью частого поиска маршрутов.

При росте скорости до 20 м/с протокол OLSR нередко оказывается эффективнее, так как топология обновляется проактивно, и при условии, что сеть не слишком велика, маршруты успевают перестраиваться заранее. Протокол AODV может демонстрировать значительные задержки, когда происходит внезапная потеря пути и нужно отправлять сообщение RREQ.

При увеличении числа узлов до 30 протокол DSDV начинает «проседать» сильнее всего: 0,13–0,18 с, тогда как протоколы AODV и OLSR показывают несколько лучшие результаты (0,08–0,15 с). Основная причина – громоздкие таблицы протокола DSDV и необходимость постоянного их обновления, что замедляет его реакцию.

Таким образом, в контексте сети FANET, где узлы могут двигаться быстро, протокол OLSR дает преимущество в плане задержки при средних скоростях, но при очень высокой скорости либо большом числе узлов ситуация может поменяться.

Другая важная метрика – пропускная способность (Throughput), особенно актуальная при использовании TCP-трафика (BulkSend), где протокол может «заполнять» всю доступную полосу. Ниже в таблице 2 приведены усредненные данные на основе нескольких экспериментов.

Таблица 2. Пропускная способность (Throughput)

Число узлов БСОС	Скорость БЛА (м/с)	DSDV	AODV	OLSR
10	5	0,4244	0,4422	0,4693
10	10	0,4318	0,4087	0,4944
10	20	0,4211	0,4410	0,4193
20	5	0,3740	0,4203	0,4455
20	10	0,4445	0,3504	0,4477
20	20	0,4147	0,3497	0,3614
30	5	0,2890	0,4261	0,4104
30	10	0,3095	0,2954	0,3413
30	20	0,2377	0,2425	0,3483

Протокол OLSR демонстрирует наибольшую пропускную способность, особенно при малом количестве узлов в беспроводной сети, что обусловлено его проактивной природой маршрутизации. Протокол AODV показывает умеренно высокие результаты, которые снижаются при увеличении скорости узлов. Протокол DSDV имеет самые низкие показатели пропускной способности, особенно при высокой подвижности сети, что связано с частыми обновлениями таблиц маршрутов.

Для 10 узлов протокол OLSR часто показывает высокие значения пропускной способности (до ~0,49 Мбит/с). Это объясняется тем, что проактивно поддерживаемая топология позволяет TCP-соединению быстрее наращивать окно congestioncontrol.

Протокол AODV при 20 узлах и скоростях БЛА 15–20 м/с нередко показывает снижение пропускной способности до ~0,35 Мбит/с, что свидетельствует о частых разрывах маршрута и необходимости заново находить путь.

Протокол DSDV зачастую оказывается аутсайдером, особенно когда число узлов достигает 30. Тогда он затрачивает больше усилий на обновление таблиц маршрутов, и TCP просто не успевает «раскачаться» до высоких скоростей передачи.

Следует отметить, что абсолютные цифры Throughput здесь могут казаться низкими по сравнению с теоретическими 11 Мбит/с Wi-Fi 802.11b, но учитывая многопрыжковую передачу, коллизии в общей беспроводной среде, накладные расходы маршрутизации, а также учет потерь при высокой мобильности – подобные величины вполне реалистичны для симуляции реальной сети FANET.

Наконец, важнейшая метрика – накладные расходы на маршрутизацию RoutingOverhead, т. е. доля или абсолютное число служебных пакетов, передаваемых для маршрутизации. Избыточный Overhead способен «забить» беспроводной канал и привести к увеличению задержек и падению значения PDR.

Таблица 3. Накладные расходы на маршрутизацию (RoutingOverhead)

Число узлов БСОС	Скорость БЛА (м/с)	DSDV	AODV	OLSR
10	5	1,1347	0,9796	1,5277
10	10	1,1687	1,1880	1,3040

Число узлов БСОС	Скорость БЛА (м/с)	DSDV	AODV	OLSR
10	20	1,8363	1,4148	1,4396
20	5	3,1511	1,6421	3,0526
20	10	3,0012	1,8516	3,3786
20	20	2,8006	2,9819	3,6373
30	5	3,5565	3,3358	3,8807
30	10	5,0567	2,1852	5,4844
30	20	6,1762	2,5567	4,0946

Накладные расходы на маршрутизацию демонстрируют значительные различия между протоколами. Протокол OLSR показывает наиболее высокие значения накладных расходов, особенно при увеличении скорости узлов, так как он использует периодические обновления таблиц маршрутизации. Протокол DSDV также имеет высокие накладные расходы, поскольку требует регулярного обновления маршрутов, особенно при увеличении числа узлов БСОС. Протокол AODV демонстрирует умеренные накладные расходы, поскольку формирует маршруты по запросу, а не поддерживает их постоянно. Протокол OLSR имеет наибольшие накладные расходы среди протестированных протоколов, особенно при высокой скорости узлов, что связано с постоянной передачей обновлений маршрутов. Протокол DSDV также показывает значительное увеличение накладных расходов с ростом числа узлов и скорости движения, что обусловлено его периодической пересылкой информации о маршрутах. Протокол AODV демонстрирует наименьшие накладные расходы благодаря использованию маршрутизации по запросу, что делает его более эффективным при высокой подвижности сети.

Протокол OLSR при 10 узлах может быть весьма эффективным, но при 20–30 узлах он начинает генерировать все больше и больше топологических сообщений, в результате чего Overhead подскакивает. Протоколу AODV при повышении скорости или числа узлов приходится чаще высылать сообщение RREQ, но всеже у него нет постоянных широковещательных рассылок всех таблиц, поэтому Overhead чаще остается в пределах 1,4–3,3 (в исключениях – до 2,98 при скорости БЛА 20 м/с и 20 узлах БСОС).

В протоколе DSDV Overhead нарастает достаточно быстро при увеличении количества узлов, так как нужно периодически передавать инкрементальные и/или полные обновления таблиц маршрутов. При 30 узлах и высокой скорости БЛА цифры могут достигать 5–6.

Исходя из данных результатов, можно найти компромиссы: более быстрое обновление маршрутов (OLSR) в высокодинамичных условиях повышает Overhead, а протокол DSDV при большом числе узлов и высоких скоростях БЛА вообще становится очень затратным. Протокол AODV – компромиссный вариант, когда не нужно постоянно поддерживать топологию, а маршруты ищутся «по требованию».

Заключение

Чтобы понять, какой протокол лучше подходит для конкретных условий работы сети FANET, надо учитывать параметры миссии:

Количество БЛА: если планируется 10–20 аппаратов, система не будет слишком велика. В таком случае протокол OLSR может обеспечить низкие задержки и высокую надёжность. При 30 и более узлах БСОС и быстрой смене топологии протокол OLSR генерирует значительный Overhead, однако все еще бывает предпочтительным, если задержка – критический фактор (например, передача видеопотока в реальном времени).

Скорость: при относительно низкой скорости (до 10 м/с) протокол DSDV не слишком страдает, но если скорость растет до 15–20 м/с, протоколы AODV или OLSR чаще будут показывать лучшие результаты.

Тип нагрузки: при TCP важно поддерживать устойчивость маршрутов для наращивания окна; при UDP (особенно CBR-трафике) критична потеря пакетов, а задержки могут варьироваться.

Из полученных данных можно сделать вывод, что для большинства сценариев с высокой подвижностью и средней/большой размерностью сети протокол AODV становится более сбалансированным выбором, позволяя поддерживать умеренные задержки. Однако если речь

идет о меньшем числе узлов (10–15) и необходимости минимизации задержки (например, для полубеспилотных миссий, где требуется время реакции), протокол OLSR может быть предпочтительнее. Протокол DSDV целесообразен только при простых условиях – медленное перемещение БЛА, небольшой рой.

Это дает основание разработчикам прикладных решений для сетей FANET выбирать соответствующий протокол:

1. При необходимости минимальной задержки и небольшом рое (до 20 узлов) есть смысл ориентироваться на протокол OLSR.
2. При быстрой смене топологии и необходимости компромисса между параметрами Overhead и PDR протокол AODV оказывается предпочтительнее.
3. Протокол DSDV может быть использован в тех случаях, когда рой БЛА очень мал (до 10 узлов) и скорость перемещения невысока, либо когда простота реализации важнее, чем оптимальная производительность.

Результаты исследования вполне коррелируют с результатами других авторов (в литературе традиционно отмечают, что протокол AODV является более гибким для динамичных MANET/FANET, тогда как протокол OLSR выгоден при стабильной/умеренно мобильной сети). Практическое значение: выводы могут быть использованы при проектировании реальных сетей БЛА для решения таких задач, как картографирование, патрулирование, мониторинг местности. В завершение следует подчеркнуть, что результаты, полученные в имитационной среде, не заменяют полностью натурные эксперименты, но дают хорошую основу для понимания ключевых закономерностей и выбора наиболее подходящего протокола на этапе проектирования.

Литература

1. Srivastava A., & Prakash J. Future FANET with application and enabling techniques: Anatomization and sustainability issues // Computer Science Review. 2021. vol. 39. p. 1059.
2. Weingartner E., vom Lehn H., & Wehrle K. A Performance Comparison of Recent Network Simulators // IEEE International Conference on Communications. 2009. pp. 1–5.
3. Tazaki H., et al. Direct Code Execution: Revisiting Library OS Architecture for Reproducible Network Experiments // Proceedings of the Ninth ACM Conference on Emerging Networking Experiments and Technologies. 2013. pp. 217–228.
4. Wang L., & Zhang Q. A Game-Theoretic Approach to Resource Allocation in FANETs // IEEE Transactionson Wireless Communications. 2022. vol. 21. no. 5. pp. 3456–3467.
5. Wang S., Zhao Y., & Li X. Security Challenges and Solutions in FANETs: A Comprehensive Survey // IEEE Communications Surveys & Tutorials. 2023. vol. 25. no. 1. pp. 789–812.
6. Zhang Y., & Liu Z. Delay-Tolerant Routing Protocol for FANETs in Urban Environments // Wireless Personal Communications. 2020. vol. 113. pp. 789–804.
7. Wang Y., & Liu J. A Mobility Model for FANETs Based on Realistic Flight Patterns // Journal of Communications and Networks. 2021. vol. 23. no. 3. pp. 210–220.

STUDY OF THE ROUTING MODEL OF INFORMATION INTERACTION CHANNELS IN THE FANET NETWORK

A.L. Priorov, A.V. Idiatullin, O.D. Moroz

P.G. Demidov Yaroslavl State University

The NS-3 software package was used to study and compare routing protocols in the wireless self-organizing (BSOS) FANET network. The OLSR, AODV, and DSDV protocols were used to implement the simulation task. The modeling parameters (number of nodes, speeds, area size, type of traffic) are defined, and the metrics used to evaluate performance (Delay, Throughput, Overhead) are described. The algorithm of the experiment has been refined, including the collection of logs, pcap analysis and the use of Flow Monitor for automatic statistics.

О НОВЫХ ПОДКЛАССАХ НЕЙРОННЫХ СЕТЕЙ ХОПФИЛДА, РЕАЛИЗОВАННЫХ НА ТЕХНОЛОГИЯХ ОПТИМИЗАЦИОННОЙ ТЕОРИИ

д.т.н., проф. Золотарев В.В.[1], д.т.н., доц. Овечкин Г.В.[2]

[1] Институт космических исследований РАН
[2] Рязанский государственный радиотехнический университет

Рассмотрены характеристики ранее не описанных где-либо подклассов нейронных сетей с обратными связями, известных как сети Хопфилда (СХ), реализованные как алгоритмы декодирования помехоустойчивых кодов и созданные на базе технологий Оптимизационной теории (ОТ) помехоустойчивого кодирования. Обсуждаются характеристики таких подклассов СХ для различных типов каналов, формирующих искажения при передаче дискретных данных. Подчёркивается правильность ориентации технологий ОТ на методы, улучшающие сходимость решений декодеров этого типа к оптимальным при большом уровне шума. Делается вывод о единстве и полном соответствии свойств и возможностей такого важного класса сетей как СХ и теории ОТ в дальнейшем развитии ключевых технологий цифрового мира. Предложены дальнейшие направления развития методов ОТ, которые могут повысить помехоустойчивость передачи и хранения данных.

Нейронные сети Хопфилда (СХ) относятся к однородным сетям с обратной связью [1]. Они первоначально предназначались, в частности, для классификации (распознавания) различных объектов в условиях некоторой доли ошибок в их исходном представлении. Эти искажения в случае их весьма ограниченного количества убирались сетью, которая могла работать в режиме последовательной (асинхронной) или одновременной (синхронной) сходимости к нескольким возможным устойчивым состояниям, неизменность которых с некоторого момента времени свидетельствовала о завершении процесса распознавания. Пример трёхнейронной СХ представлен на рисунке 1.

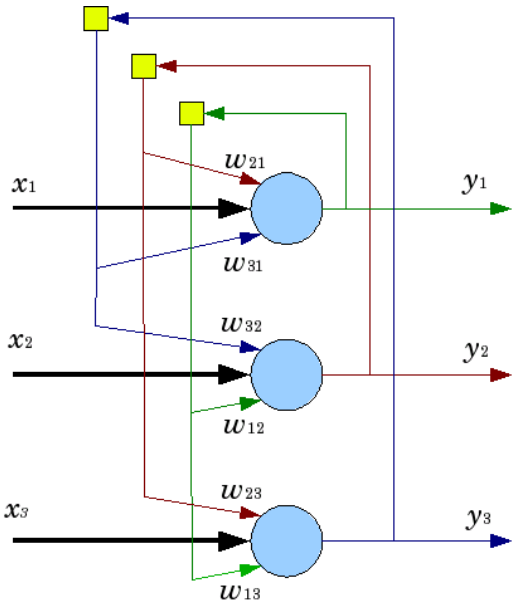


Рисунок 1 – Трёхнейронная сеть Хопфилда

Первоначально СХ считается однослойной, а нейроны двоичными, т.е. имеющими 2 возможных состояния. Сеть в процессе распознавания осуществляет вычисления, которые при каждом изменении её состояния минимизируют функционал, обычно называемый энергией сети. В процесс обучения сети входит и выбор этого функционала, т. е. правила выполнения таких вычислений в сети. Хорошо известно, что число возможных различных объектов, распознаваемых сетью СХ на порядок меньше её размера. А вопросы соотношения числа таких объектов и количества ошибок в первоначально загружаемых в сеть их образах малоисследованы [1, 2].

Мажоритарные декодеры, в частности, многопороговые алгоритмы декодирования (МПД), созданные в процессе разработки Оптимизационной Теории (ОТ) помехоустойчивого кодирования также относятся к сетям типа СХ.

Пример МПД для блокового двоичного кода длины n=26 с k=13 информационными символами и кодовой скоростью R=1/2 представлен ниже на рисунке 2 [3]. Характеристики МПД разных типов детально рассмотрены и изучены в [4-6].

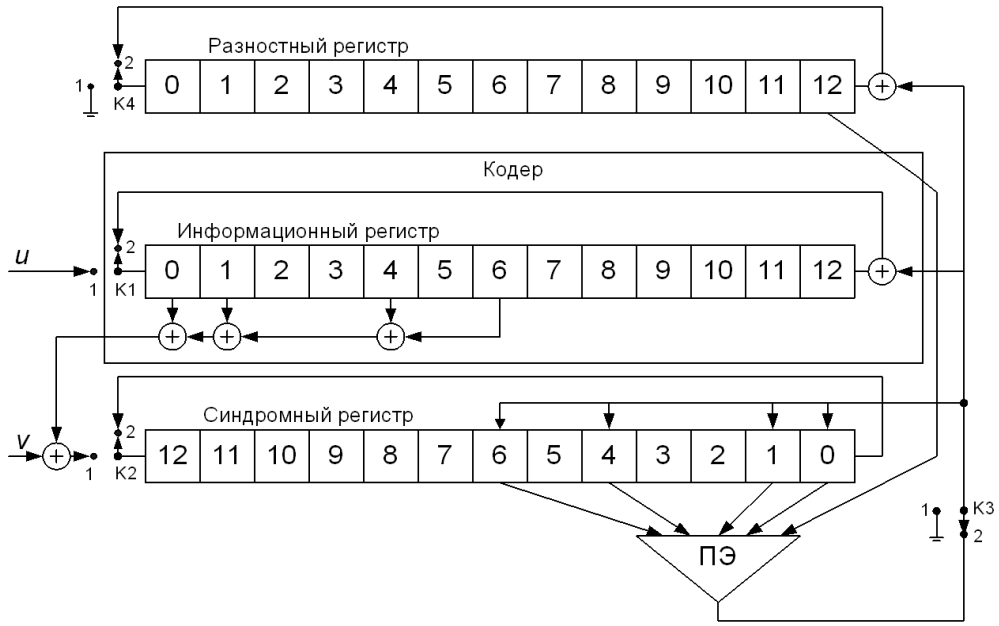


Рисунок 2 – МПД декодер как один из вариантов СХ

Данный декодер получен из схемы на рисунке 1 разделением функций, выполняемых разными нейронами, и минимизацией числа активных вычисляющих нейронов, что значительно уменьшает реальную аппаратную сложность всей СХ в целом в режиме асинхронной работы, что всегда и получается при программном моделировании СХ. В терминах СХ энергия такой схемы из двоичных нейронов равна общему числу «1» в разностном и синдромном регистрах, а вычисляемая функция, естественно, пороговая, равная «1», если обычная сумма на пороговом элементе (ПЭ), который принимает решение о значении информационного символа в нейроне 12, превышает половину общего числа входов для этого ПЭ. Тогда через обратные связи значения всех нейронов, с которых поступали данные на ПЭ, инвертируются. Таким образом, при изменении информационных символов всегда энергия уменьшается не менее, чем на 1, все требования к схеме СХ выполняются, и она оказывается устойчивой, так как в какой-то момент времени состояние сети перестанет изменяться, поскольку число «1» в указанных двух регистрах станет маленьким и пороговое значение сумм на входах ПЭ уже не будет превышаться. Тем самым СХ достигнет одного из возможных локальных или глобального минимума энергии.

А с точки зрения теории кодирования (ТК) схема МПД декодера при каждом изменении декодируемых символов и соответствующих ячеек в синдромном и разностном векторах находит всё более правдоподобное решение. Но декодер при этом может остановиться, как достигнув, так и не «добежав» до истинного оптимального (наилучшего по достоверности) решения. Все эти эффекты неполной сходимости в теории ОТ давно и детально хорошо изучены. На наличие этих особенностей сходимости указывается и во многих публикациях по общей теории для СХ.

А далее обратим внимание на то, что в теории кодирования (ТК) в рассматриваемой схеме декодирования сам алгоритм МПД уже не различает конфигурации нескольких объектов при небольшом числе искажений в их первоначальном представлении, а выполняет фильтрацию ошибок в единственном объекте, причём, только в части нейронов, находящихся в нейронах информационного регистра декодера. Но при этом в полном согласии с известными в ТК границами для исправляемой доли ошибок МПД алгоритм исправляет уже очень значительное

обычно число первоначальных ошибок, помещаемых в такой «МПД-СХ». Именно дальнейшее рассмотрение СХ как МПД декодера, созданного в соответствии с теорией ОТ, как мы увидим далее, создаёт условия для исследования и анализа целого «ансамбля» принципиально новых модификаций СХ, которые, как мы полагаем, ещё нигде ранее не рассматривались.

Обратимся для этого к графикам на рисунке 3, который взят из рабочих лабораторных журналов, отслеживающих эксперименты с МПД декодерами. Кривая *1* соответствует зависимости вероятности ошибки декодирования $P(e)$ блокового кода длины n=250 от вероятности P_0 случайно вносимых в сеть ошибок извне, т.е. при её загрузке. Это типичная задача определения эффективности декодирования, например, в канале типа двоичного симметричного (ДСК) [5, 7]. Здесь и далее везде R=1/2, а кодовое расстояние d используемых самоортогональных кодов (СОК) [4] равно 11. Существенно, что разрешённое количество попыток исправления, т.е. итераций декодирования содержимого тех информационных битов, которые всегда только и нужны далее пользователю такой «нейросети-декодера», равно I=10. В различных реальных кодовых системах величина I может варьироваться от 3 до ~200.

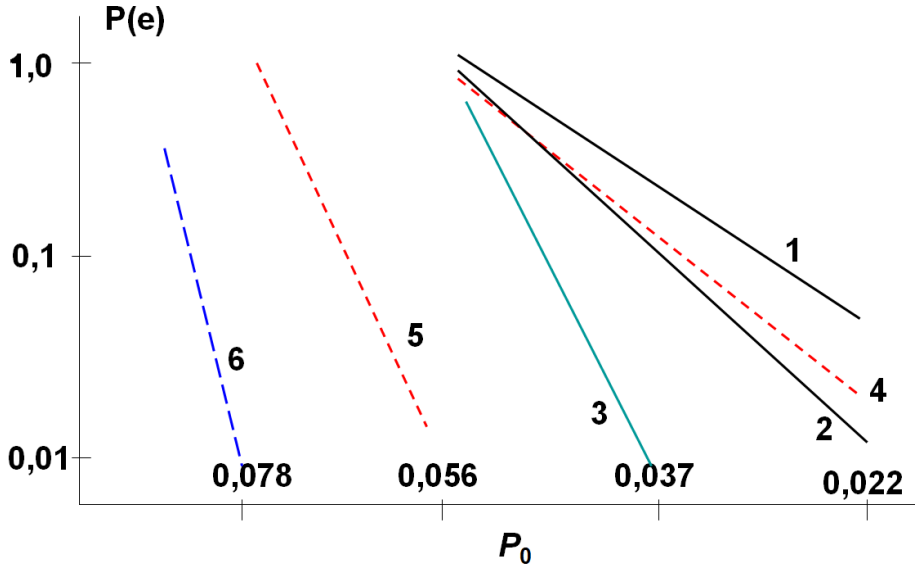


Рисунок 3 – Характеристики новых подклассов нейросетей Хопфилда

А далее будут представлены варианты систем «МПД-СХ» с совершенно новыми параметрами, возможностями и характеристиками, о которых мы не видели никаких близких к нам публикаций по тематике СХ. Кривая *2* соответствует случаю, когда в отличие от базового варианта СХ выполняемые функции на каждой итерации несколько варьируются. В данном конкретном случае применена оптимизация нескольких десятков параметров СХ на разных итерациях декодирования. При этом сохраняется правило, согласно которому все решающие функции на каждой итерации уменьшают, как и в базовой сети, её энергию. Как видим, характеристики системы улучшились, т.е. вероятность ошибочного конечного состояния снизилась.

Следующий график *3* относится уже к длинному коду с n=4000, но с теми же значениями R=1/2 и d=11, как у предыдущих графиков. Для этого декодера были сразу оптимизированы по технологиям ОТ некоторые из параметров МПД. Существенно, что длина очень сильно улучшает характеристики декодирования созданной системы «МПД-СХ», как это и следует из самых основ ТК и декодер успешно работает при более высоком уровне шума канала на его входе. В [5, 7, 8] показано, что используемые в МПД коды могут подбираться по очень многим параметрам, которые нередко улучшают работу декодера очень существенно и, что важно, весьма простыми средствами.

Графики *4* и *5* относятся к давно известным МПД алгоритмам для гауссовских каналов. Но при этом система «МПД-СХ» уже не может оставаться двоичной, так как наличие квантования сигнала на входе декодера заставляет его работать с небольшими целыми числами, которые относятся к надёжности приёма декодируемого бита в МПД. При этом энергия нейросети, как и раньше, благодаря корректировке решающих функций на всех этапах (итерациях), уменьшается

при всех изменениях информационных символов кода. Но в явном виде такой СХ мы в текущей литературе не видели. МПД декодеры для гауссовского канала известны с 1976 года и описаны в справочниках, опубликованных научной школой ОТ [5, 8]. Они доступны на наших сетевых порталах [6]. Отметим в связи с этим, что собственно Оптимизационная Теория завершена ещё в 1985 году и, по мнению ряда членов и академиков РАН, полностью разрешила великую проблему Шеннона от 1948 г. о простой высокодостоверной передаче дискретной информации по сильно шумящим цифровым каналам. Это значит, что все декодеры, созданные по технологиям ОТ имеют линейную по длине кодов n сложность и при этом даже при весьма высоком относительном уровне шума цифрового канала практически всегда обеспечивают наилучшие по достоверности оптимальные решения. Успех школы ОТ в решении великой проблемы Шеннона отмечен в [9], где сторонников школы ОТ коллеги поздравили с тремя юбилеями ТК, к которым школа ОТ имеет непосредственное отношение.

Следующий новый подкласс систем «МПД-СХ» образуют символьные коды, как назвали особые недвоичные коды с мажоритарным декодированием и простейшим декодированием. Детально основы теории кодирования для символьных кодов можно посмотреть в [5-8]. Мы ограничимся пока одним примером символьного МПД со скоростью R=1/2, d=11 длиной n=1000 и I=10, обладающего всеми необходимыми свойствами СХ. Его график для размера алфавита недвоичного кода q=256 представлен кривой *6*. Более длинные коды в символьном МПД при достаточно большом числе итераций I могут, как показано в [3, 5-7], успешно работать при вероятностях ошибок на входе порядка 0,25 и даже более.

Наконец, ограничимся замечанием о том, что в каналах со стираниями соответствующие МПД также удовлетворяют требованиям к сетям СХ по виду решающей функции. Эти МПД успешно работают, заменяя почти все стёртые символы их правильными значениями, даже при очень высоких вероятностях стираний в канале, которые при R=1/2 могут быть даже больше, чем 0,4. Этим примером мы и завершим обзор полного соответствия всех типов МПД декодеров основным обязательным принципам работы СХ.

Отметим, что наиболее эффективными в высокоскоростных каналах спутниковых связи нередко оказываются разнообразные непрерывные коды, чаще всего называемые специалистами свёрточными. Развитие теории и технологий ОТ начиналось в 70-х годах прошлого века тоже со свёрточных кодов, которые оказываются в целом ряде приложений предпочтительнее блоковых. Свёрточные МПД поэтому стали известны ещё раньше блоковых и также удовлетворяют всем принципам, описывающим структуру и работу СХ.

Вся единая теория ОТ для блоковых и свёрточных кодов изложена в [5, 7], а также в других работах школы ОТ. Возможно, что ранее СХ, соответствующие бесконечным кодовым последовательностям, также не описывались где-либо. В этом случае СХ для «бесконечных», т.е. очень больших массивов данных со свёрточной структурой также можно отнести на счёт достижений школы ОТ.

Общий вывод по проведённым пятидесятилетним исследованиям: Опыт конкретных реализаций СХ в виде технологий ОТ в важной и очень сложной сфере борьбы с искажениями данных при передаче по шумящим каналам оказался много шире и разнообразнее, чем на текущий момент это сделано для СХ достаточно общего типа. Так и должно быть во всех конкретных глубоких разработках, учитывающих специальные условия работы таких сетей СХ во всех возможных их приложениях.

Успешность и теперь уже абсолютное лидерство теории и технологий ОТ в полном разрешении проблемы Шеннона отлично подтверждено её единством со многими уже известными нейросетевыми подходами к решению проблемы искажений, а также глубоким осмыслением конкретных процедур ОТ и постоянным поиском новых возможностей улучшения характеристик ОТ средствами её собственных технологий, находящихся в прекрасном соответствии с реальными ресурсами как цифрового, так и аналогового мира.

Литература

1. Нейронная сеть Хопфилда – Википедия.
2. Уоссермен Ф. Нейрокомпьютерная техника: Теория и практика. М.: Мир, 1992. 240 с.

3. Золотарёв В.В., Овечкин Г.В. и др. Проблемы и открытия Оптимизационной Теории (ОТ в иллюстрациях) // Под редакцией академика РАН Н.А. Кузнецова. М.: ИКИ РАН, «Горячая линия – Телеком», 2020, 38 с., https://decmtdzol.ru/articles/comics.pdf .
4. Золотарёв В.В., Овечкин Г.В. Помехоустойчивое кодирование. Методы и алгоритмы. Справочник // Под научной редакцией члена-корреспондента РАН Ю.Б. Зубарева. М.: «Горячая линия – Телеком», 2004, 126 с.
5. Золотарёв В.В., Зубарев Ю.Б., Овечкин Г.В. Многопороговые декодеры и Оптимизационная Теория кодирования // Под редакцией академика РАН В.К. Левина. М.: «Горячая линия – Телеком», 2012, 238 с. https://decmtdzol.ru/articles/Zolotarev_OK2012.pdf.
6. Сетевые ресурсы научной школы ОТ: www.mtdbest.ru, www.decoders-zolotarev.ru, www.decmtdzol.ru.
7. Золотарёв В.В. Кодирование для цифровой связи и систем памяти. Справочник-2. Под научной редакцией члена-корреспондента РАН Ю.Б. Зубарева. М.: «Горячая Линия – Телеком», 2022, 172 с. https://mtdbest.ru/books_rus.php.
8. Zolotarev V.V., Zubarev Y.B., Ovechkin G.V. Optimization Coding Theory and Multithreshold Algorithms. Geneva, ITU, 2015, 159 p. https://mtdbest.ru/articles/Zolotarev_ITU.pdf.
9. Овечкин Г.В., Пылькин А.Н., О трех значимых юбилеях теории кодирования // Вестник РГРТУ. 2024. №90. https://mtdbest.ru/articles/Anniversary_vestnik2024.pdf.

ON NEW SUBCLASSES OF HOPFIELD NEURAL NETWORKS IMPLEMENTED USING OPTIMIZATION THEORY TECHNOLOGIES

Zolotarev V.V.[1], Ovechkin G.V.[2]

[1] Space Research Institute RAS
[2] Ryazan State Radio Engineering University

The characteristics of previously undescribed subclasses of neural networks with feedback, known as Hopfield networks (HN), implemented as algorithms for decoding error-correcting codes and created on the basis of the technologies of the Optimization Theory (OT) of error-correcting coding are considered. The characteristics of such subclasses of HN for various types of channels that generate distortions during transmission of discrete data are discussed. The correctness of the orientation of OT technologies on methods that improve the convergence of solutions of this type of decoders to optimal ones at a high noise level is emphasized. A conclusion is made about the unity and full compliance of the properties and capabilities of such an important class of networks as HN and the OT theory in the further development of key technologies of the digital world. Further directions for the development of OT methods that can improve the noise immunity of data transmission and storage are proposed.

ПРИМЕНЕНИЕ КОДОВ С НЕРАВНОЙ ЗАЩИТОЙ СИМВОЛОВ ДЛЯ ПЕРЕДАЧИ ИЗОБРАЖЕНИЙ

д.т.н., проф. Золотарев В.В.[1], д.т.н., доц. Овечкин Г.В.[2], к.т.н., доц. Гринченко Н.Н.[2]

[1] Институт космических исследований РАН
[2] Рязанский государственный радиотехнический университет

Рассматривается задача передачи полутоновых изображений по каналам с шумом. Выполнено исследование эффективности известных методов помехоустойчивого кодирования при передаче изображений по каналу с аддитивным белым гауссовским шумом. Показано, что применение многопороговых декодеров самоортогональных кодов (СОК) в таких условиях позволяет эффективно решать данную задачу, обеспечивая сопоставимую с LDPC кодами эффективность при меньшей вычислительной сложности. Предложена новая схема адаптивного кодирования изображений на основе СОК с неравной защитой символов, обеспечивающая на 0,5 дБ лучшие характеристики по сравнению с LDPC кодами.

Введение

В последние десятилетия все большее значение имеет технология дистанционного зондирования Земли, в процессе применения которой с помощью датчиков (камеры, сканеры, радары и пр.), установленных на различные летательные аппараты, получаются данные об объектах на поверхности земли без физического контакта с ними. Эти данные, например, в виде аэрофотоснимков или спутниковых снимков, должны быть переданы на землю для дальнейшей обработки. При этом, безусловно, необходимо использовать современные технологии передачи цифровых данных, обеспечивающие эффективное использование очень дорогих ресурсов канала связи [1]. Одной из таких технологий является помехоустойчивое кодирование, применяемое для исправления ошибок при передаче данных по каналам с шумом. Следует отметить, что на сегодняшний день известен целый ряд помехоустойчивых кодов и достаточно эффективных методов их декодирования, отличающихся корректирующей способностью, вносимой избыточностью, длиной, сложностью применяемых алгоритмов и многими другими параметрами [1, 2]. В работе выполнен анализ таких методов коррекции ошибок для передачи изображений и предлагаются варианты улучшения их эффективности.

Особенности задачи помехоустойчивого кодирования при передаче изображений

В работе ограничим круг передаваемых изображений полутоновыми изображениями с 256 уровнями оттенков серого. Такие изображения могут быть представлены матрицей размера $M \times N$:

$$\mathbf{B} = \begin{bmatrix} b_{1,1} & b_{1,2} & \dots & b_{1,N} \\ b_{2,1} & b_{2,2} & \dots & b_{2,N} \\ \dots & \dots & \dots & \dots \\ b_{M,1} & b_{M,2} & \dots & b_{M,N} \end{bmatrix}.$$

Следует отметить, что представленные далее результаты можно распространить и на большинство других типов изображений.

В процессе передаче таких изображений выполняются этапы помехоустойчивого кодирования, модуляции, передачи по каналу, демодуляции и декодирования, в результате чего на приемной стороне получается принятое изображение $\hat{\mathbf{B}}$, которое может отличаться от переданного из-за возможных неисправленных канальных ошибок. При этом, в отличие от классической задачи передачи обобщенных цифровых данных, в которой качество может оцениваться в терминах частоты ошибок на бит (BER) или блок (BLER), в задаче передачи изображений качество целесообразно

оценивать в терминах среднего квадратического отклонения (СКО, MSE) между яркостями пикселов исходного и принятого изображений как хорошо согласующееся с физиологическими особенностями зрения человека. СКО между яркостями $b_{i,j}$ и $\hat{b}_{i,j}$ пикселей переданного и принятого изображения можно вычислить по формуле

$$\sigma = \sqrt{\frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}\left(\hat{b}_{i,j} - b_{i,j}\right)^2} .$$

Известно, что изображение с уровнем СКО ниже 2,5 неотличимо от оригинала, с уровнем от 2,5 до 4,5 показывает хорошее качество, с уровнем выше от 4,5 до 8 показывает удовлетворительное качество, выше 8 – неудовлетворительное. Из этого следует, что при передаче изображений нет необходимости применения сверхсложных методов помехоустойчивого кодирования, обеспечивающих очень малые вероятности ошибки, например на уровне 10^{-9} и менее, а достаточно при как можно большем уровне шума обеспечить приемлемое для дальнейшего использования изображений СКО, не превышающее 2,5. Так же можно отметить, что большое количество получаемых изображений и их высокое разрешение приводят к необходимости передачи огромного объема данных, что требует минимизировать сложность алгоритмов помехоустойчивого кодирования и декодирования.

Анализ эффективности методов коррекции ошибок при передаче изображений

Рассмотрим эффективность известных методов помехоустойчивого кодирования при передаче изображений, полученную с помощью компьютерного моделирования. При этом будем использовать наиболее простую модель канала с аддитивным белым гауссовским шумом, двоичную фазовую модуляцию и демодулятор, принимающий жесткие решения относительно принятых символов. Такая модель хорошо подходит, например, для спутниковых каналов связи. В качестве кодов будут использоваться блоковые коды с кодовой скоростью $R \approx 4/5$.

На рисунке 1 представлена зависимость СКО между переданным и принятым изображением от битового отношения сигнал-шум в канале связи. Кривая *1* соответствует эффективности кода Рида-Соломона (РС) (255, 205, 41). Это символьный код, способный исправить до 20 байтовых ошибок в кодовом блоке. Видно, что с помощью такого кода можно обеспечить СКО≤2,5 начиная с E_b/N_0=5,5 дБ. При этом без применения кодирования такое СКО обеспечивалось при E_b/N_0=7,8 дБ, т.е. выигрыш в энергетике составил около 2,3 дБ. Также отметим, что пропускная способность, при которой можно обеспечить сколь угодно малую вероятность ошибки, для кодовой скорости R=4/5 составляет E_b/N_0=3,37 дБ, т.е. теоретически при такой избыточности можно работать еще на 2,1 дБ большем уровне шума, чем для кодов РС. При использовании значительно более эффективных низкоплотностных (LDPC) кодов стандарта DVB-S2 длиной 64800 битов и аналогичной кодовой скоростью можно обеспечить требуемое СКО при $E_b/N_0 \approx 4{,}4$ дБ, что на 1 дБ лучше, чем для кодов РС. Вместе с тем кодер и декодер для данных кодов являются не самыми простыми с вычислительной точки зрения и требуют больших объемов памяти при аппаратной реализации.

Поэтому для повышения эффективности системы передачи изображений целесообразно применить коды более эффективные, чем коды Рида-Соломона, имеющие малую вычислительную сложность как кодирования, так и декодирования. К таким кодам можно отнести самоортогональные коды (СОК), для которых существуют эффективные многопороговые алгоритмы декодирования (МПД) [3, 4, 5]. При этом СОК могут быть блоковыми и сверточными, двоичными и символьными, короткими и

длинными и иметь различную кодовую скорость. Это дает возможность гибкой подстройки применяемой системы кодирования под решаемые задачи.

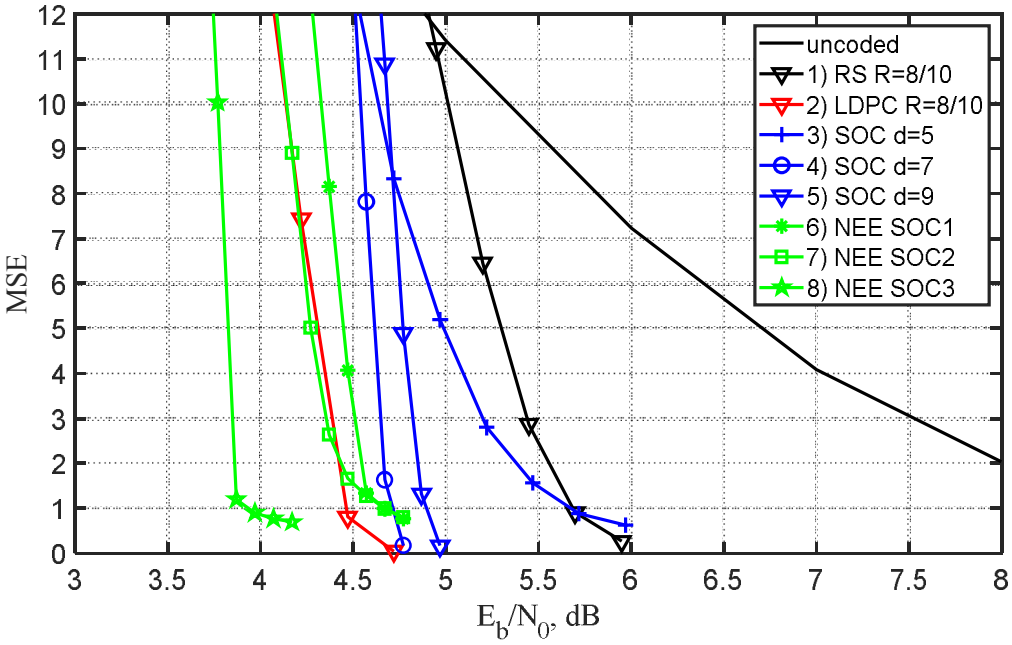


Рисунок 1 – СКО изображения при использовании различных помехоустойчивых кодов с R=4/5

На рисунке 1 представлены зависимости СКО между переданным и принятым изображением при использовании двоичных блоковых СОК с кодовой скоростью 4/5. При этом кривая *3* соответствует коду с кодовым расстоянием 5, кривая *4* – с кодовым расстоянием 7, а кривая *5* – с кодовым расстоянием 9. Отметим, что, с одной стороны, коды с большим кодовым расстоянием обеспечивают меньшую вероятность ошибки, и поэтому способны обеспечить меньшее СКО. Но область, в которой они обеспечивают меньшее СКО, сдвигается в сторону меньших уровней шума. Это объясняется существенным увеличением размерности проверок с ростом кодового расстояния, в результате чего при большом шуме вероятность их ошибочности становится близкой к 0,5, т.е. они оказываются бесполезными. Поэтому для задач передачи изображений наилучшим будет использование СОК с кодовым расстоянием 7, позволяющими получить СКО≤2,5 начиная с отношения сигнал-шум около 4,7 дБ, что на 0,7 дБ лучше, чем при использовании кодов РС с такой же кодовой скоростью. Отметим, что это преимущество обеспечивается наряду с более эффективным многопороговым алгоритмом декодирования применением двоичных и при этом значительно более длинных кодов, чем коды РС (длина применяемых СОК около 65536 битов), что при передаче изображений, например, размером 256x256 пикселей является вполне допустимым. Так же отметим, что такие СОК проигрывают LDPC кодам всего около 0,3 дБ при почти на порядок меньшей сложности декодирования.

Повышение эффективности передачи изображений

При анализе процесса передачи изображений можно заметить, что искажение различных битов яркостей пикселей по-разному влияет на восприятие изображения: искажения старших битов заметны для глаза, в то время как искажения младших битов практически не видны. Поэтому в целях повышения эффективности системы передачи изображений предлагается старшие биты яркостей защищать более сильным помехоустойчивым кодом, а младшие – более слабым или не защищать вообще. Это позволит обеспечить требуемое СКО при большем уровне шума в канале для заданной кодовой скорости. Такой вид организации кодирования будем называть адаптивным кодированием изображений.

Основной сложностью применения такого варианта организации кодирования является необходимость применения до восьми различных кодеров и декодеров (для однобайтовых яркостей), что увеличивает сложность приемопередающей аппаратуры. Для решения данной проблемы отметим, что применяемые совместно с МПД СОК имеют по несколько информационных и несколько проверочных ветвей. И для каждой пары информационная-проверочная ветвь задается свой порождающий полином, используемый при кодировании и декодировании этого кода. Но тогда, используя полиномы разной степени между разными информационными и проверочными ветвями, можно получить коды, в которых подкоды для различных информационных ветвей будут иметь различное кодовое расстояние, а, соответственно, различные информационные символы будут по-разному защищены. При этом из-за простого перераспределения проверок с одних информационных ветвей на другие не происходит роста их размерности (ее можно даже уменьшить), и эти коды будут исправлять ошибки в защищенных информационных ветвях эффективно даже при большом уровне шума, естественно, за счет меньшего исправления ошибок в незащищенных ветвях. Такие коды, называемые далее кодами с неравной защитой символов (НЗС) [3], смогут обеспечить защиту битов в соответствии с их влиянием на качество изображения. Схема передачи изображений с кодами с НЗС показана на рисунке 2.

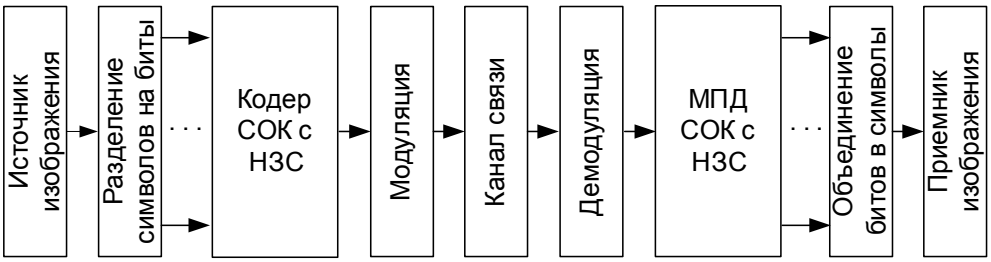


Рисунок 2 – Схема передачи изображений по каналам связи с кодами с НЗС

На рисунке 3 показан пример структуры кода с неравной защитой символов при числе информационных ветвей n_k=8, числе проверочных ветвей n_r=2, кодовой скорости $R=n_k/(n_k+n_r)=8/10=4/5$. В ячейках таблицы указано число отводов от i-й информационной на j-ю проверочную ветви. Справа от таблицы указано кодовое расстояние d_i соответствующей информационной ветви. Внизу – размерность проверочных ветвей. Изменяя число отводов с информационной на проверочные ветви можно увеличивать или уменьшать кодовое расстояние и, тем самым, управлять степенью защиты соответствующих битов. Так же отметим, что выбирая длину одной информационной ветви в соответствии с размером изображения можно выполнить его кодирование одним кодовым словом.

3	3	7
3	3	7
2	2	5
2	2	5
2	2	5
1	1	3
1	1	3
1	1	3
15	15	

Рисунок 3 – Пример структуры кода с НЗС

Пример схемы кодера сверточного СОК с НЗС для R=2/4 показан на рисунке 4. В нем можно выделить две информационные ветви, для первой из которых кодовое

расстояние равно 9, а для второй – 3. Следовательно, после декодирования такого кода с помощью соответствующего ему МПД (рисунок 5) в битах первой ветви будет исправлено значительно больше ошибок, чем во второй. При этом размерность проверочных вервей по сравнению с обычным СОК с d=9 уменьшена с 8 до 5, что позволяет декодеру эффективно работать при несколько большем уровне шума.

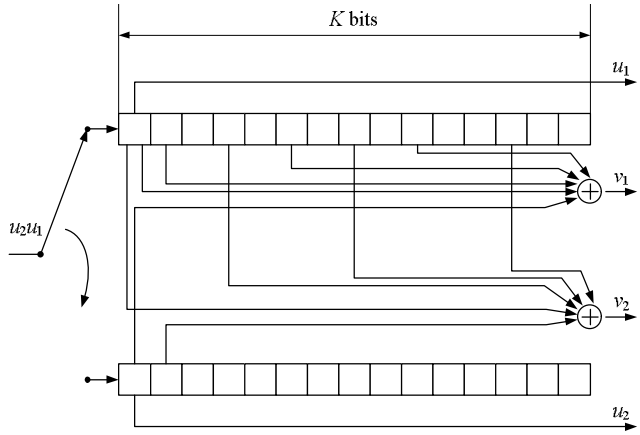


Рисунок 4 – Пример схемы кодера сверточного СОК с НЗС

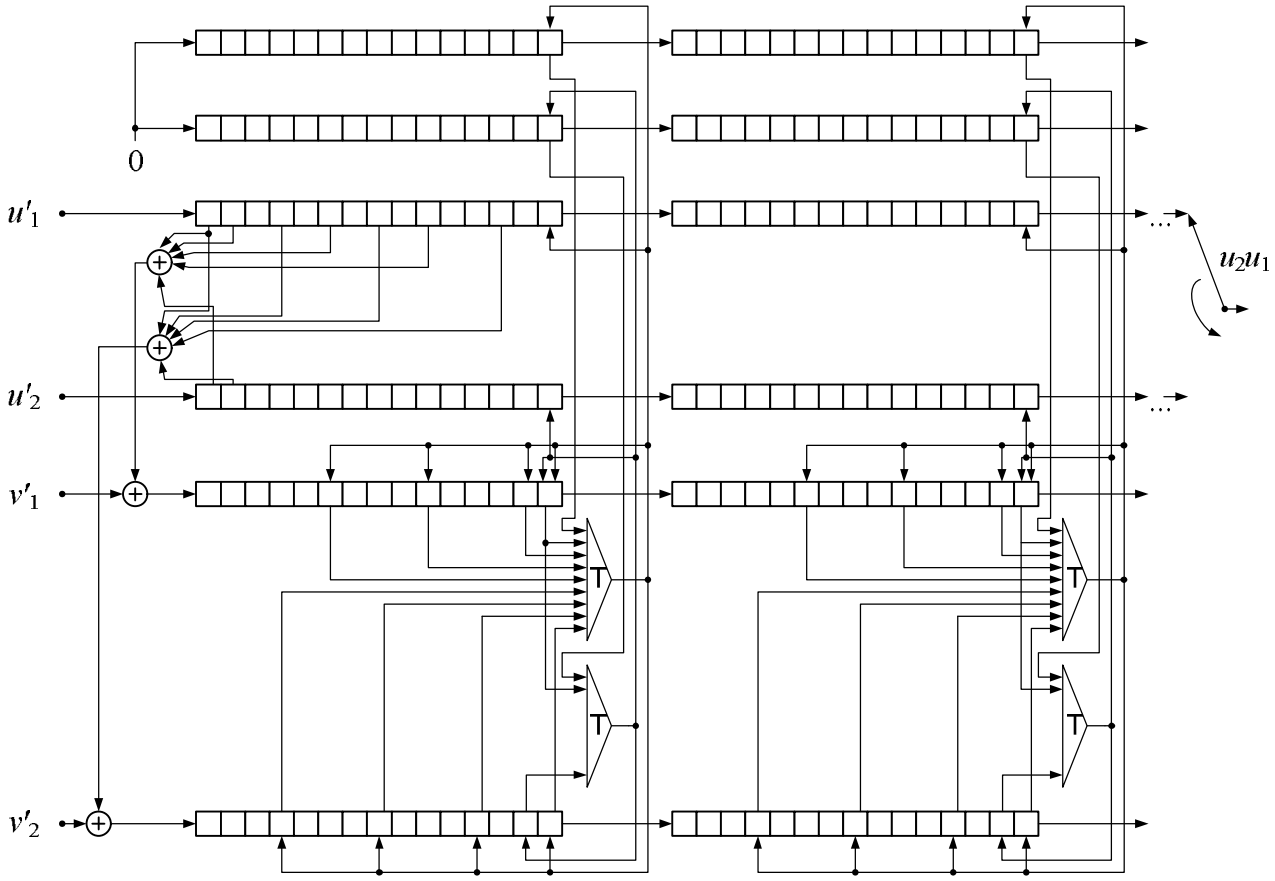


Рисунок 5 – Пример схемы МПД для сверточного СОК с НЗС

Эффективность некоторых построенных авторами СОК с НЗС с кодовой скоростью 4/5 показана на рисунке 1. Кривая *6* соответствует СОК с кодовыми расстояниями для подкодов (11, 11, 9, 9, 7, 7, 5, 5). Такой код работает немного лучше (примерно на 0,2 дБ), чем ранее рассматривавшийся СОК с d=7. Кривая *7* соответствует СОК с кодовыми расстояниями для подкодов (13, 11, 9, 7, 5, 3, 3, 3), в этом коде старшие биты еще больше защищены. Такой код обеспечивает дополнительный выигрыш еще около 0,1 дБ при СКО≤2,5, обеспечивая эффективность, сопоставимую с

эффективностью значительно более сложного для декодирования декодера LDPC кода. И, наконец, кривая *7* соответствует СОК с кодовыми расстояниями для подкодов (17, 13, 9, 5, 5, 0, 0, 0). В нем значительно усилена защита старших битов пикселей изображения и совсем не защищены три младших. Такой код обеспечивает СКО≤2,5, начиная с отношения сигнал-шум E_b/N_0≈3,8 дБ, обеспечивая дополнительный выигрыш более 0,5 дБ по сравнению с LDPC кодом и около 1,7 дБ по сравнению с кодами РС. Отметим, что длина этого кода составляет порядка полумиллиона битов, что позволяет кодировать одним кодовым блоком все изображение размером 256x256 символов.

Выводы

В работе рассмотрены подходы к решению задачи обеспечения высокого качества принятого изображения при большом уровне шума в канале, основанные на применении помехоустойчивых кодов. Предложены идеи адаптивного кодирования изображений и получены реализующие их новые самоортогональные коды с неравной защитой символов. Показано, что за счет применения таких кодов можно обеспечить СКО принятых изображений на приемлемом уровне при на 1,7 дБ большем уровне шума по сравнению с использованием кодов Рида-Соломона и на 0,5 дБ по сравнению с LDPC кодами стандарта DVB-S2, что является серьезным результатом. Этот выигрыш в энергетике канала можно использовать для улучшения важных характеристик систем передачи изображений: дальность, скорость передачи, время автономной работы передатчика, размеры антенн, ширина полосы частот и т.п.

Литература

1. Скляр Б. Цифровая связь. Теоретические основы и практическое применение // СПб., Вильямс, 2017. 1100 с.
2. Луферчик А.В., Луферчик П.В., Галеев Р.Г., Богатырев Е.В. Анализ методов помехоустойчивого кодирования в модемах станций спутниковой связи нового поколения // Успехи современной радиоэлектроники, 2023, №10. С.41-46 .
3. Zolotarev V., Zubarev Y., Ovechkin G. Optimization Coding Theory and Multithreshold Algorithms. Geneva, ITU. 2015, 159 p.
4. Zolotarev V.V., Ovechkin G.V., Zung Ch.T. The Prospects of Optimization Theory Application for Solving Shannon Problem // 2022 24th International Conference on Digital Signal Processing and its Applications, DSPA 2022.
5. Золотарёв В.В. Оптимальные алгоритмы декодирования Золотарёва. Под научн. ред. члена-корреспондента РАН Ю.Б. Зубарева // М.: Горячая линия — Телеком, 2021, 262 с.

THE USE OF CODES WITH UNEQUAL CHARACTER PROTECTION FOR IMAGE TRANSMISSION

Zolotarev V.V.[1], Ovechkin G.V.[2], Grinchenko N.N.[2]

[1] Space Research Institute RAS
[2] Ryazan State Radio Engineering University

The problem of transmitting grayscale images through noisy channels is considered. A study of the effectiveness of known error-correction codes in image transmission over a channel with additive white Gaussian noise has been performed. It is shown that the use of multithreshold decoders of self-orthogonal codes (SOC) in such conditions makes it possible to effectively solve this problem, providing efficiency comparable to LDPC codes with less computational complexity. A new adaptive image encoding scheme based on SOC with unequal character protection is proposed, providing 0.5 dB better performance compared to LDPC codes.

ИССЛЕДОВАНИЕ ЭФФЕКТИВНОСТИ АЛГОРИТМА СТАБИЛИЗАЦИИ УРОВНЯ ЛОЖНОЙ ТРЕВОГИ С ВЫБОРОМ МОДЕЛИ АППРОКСИМАЦИИ СРЕДНЕГО ЗНАЧЕНИЯ ШУМА

д.т.н., проф. Белокуров В. А.[1], асп., Нгуен Ч.К.[1]

[1]Рязанский государственный радиотехнический университет имени В.Ф. Уткина, Рязань

В данной работе приведён анализ эффективности алгоритма стабилизации уровня ложной тревоги с выбором модели аппроксимации среднего значения шума в пределах «скользящего» окна. Рассмотрен выигрыш в пороговом отношении сигнал-шум предлагаемого алгоритма по сравнению с известным алгоритмом УС-СУЛТ. Результаты анализа показали, что на фоне шума, среднее значение которого меняется в пределах «скользящего» окна предлагаемый алгоритм обеспечивает выигрыш в пороговом отношении сигнал-шум до 5,78 *дБ* по сравнению с известным алгоритмом УС-СУЛТ.

Введение

В радиолокации одной из важных задач является обеспечение стабилизации уровня ложных тревог (СУЛТ) при обнаружении целей на фоне шума. Порог обнаружения является функцией дисперсии шума. Наличие помех в «скользящем» окне, используемому для оценки дисперсии шума, может привести к изменению эффективности алгоритма СУЛТ с усреднением (УС-СУЛТ) [1].

В настоящее время в литературе представлено множество обнаружителей СУЛТ, каждый из которых показывает эффективность в тех или иных условиях наличия фона шума и помех [2], однако изменение среднего значения уровня шума в пределах «скользящего» окна не учитывается. Это приводит к ошибкам вычисления дисперсии шума и, как следствие, к изменению порога обнаружения.

В работе [3] рассмотрен алгоритм стабилизации уровня ложной тревоги при обнаружении на фоне шума, среднее значение которого меняется в пределах «скользящего» окна. Недостатком подхода является необходимость априорного знания наклона прямой, которая аппроксимирует изменение среднего значения шума. Использование аппроксимаций на основе кривых более высокого порядка в работе затруднительно.

Предлагаемый авторами адаптивный алгоритм СУЛТ реализует адаптивный выбор порога обнаружения на основе выбора модели аппроксимации среднего значения шума в пределах «скользящего» окна размером M.

Постановка задачи

На рисунке 1 показан спектр отраженного сигнала в каналах по дальности, полученный путём обработки данных РЛС с непрерывным излучением.

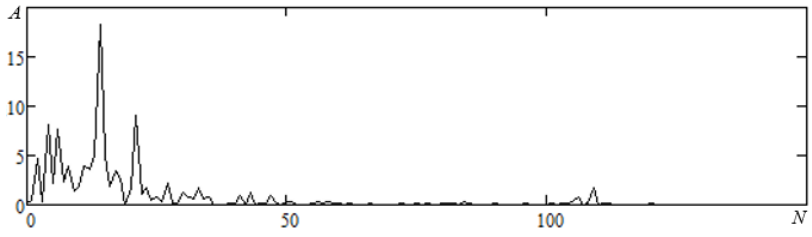


Рисунок 1 – Амплитуды отраженных импульсов в одном из каналов по дальности

На рисунке 1 введены следующие обозначения: N – число каналов по дальности; A – амплитуда.

На рисунке 2 показана обобщенная схема формирования одномерного «скользящего» окна, которое используется для оценки параметров шума.

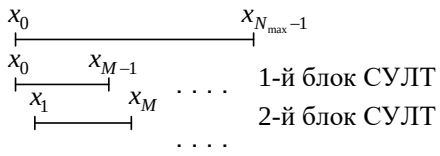


Рисунок 2 – Формирование «скользящего» окна с размером M

На рисунке 2 введены следующие обозначения блоков: $N_{\max}$ - максимальное число каналов обнаружения. На вход каждого k-го блока СУЛТ поступает M отсчётов, каждый из которых представляет собой результат вычисления достаточных статистик в соответствующих каналах обнаружения. Число блоков СУЛТ равно числу каналов обнаружения $N_{\max}$.

Введем следующие обозначения: $\mathbf{z}_k$ - массив входных отсчётов k-го блока СУЛТ, $\mathbf{z}_k = \{x_i\}_{i=k-1}^{k-1+M-1}$, $k = 1...N_{\max}$. Элементы массива $\{x_i\}_{i=0}^{N_{\max}-1}$ представляют собой квадраты модулей отсчёт на выходе накопителей [4].

Предположим, что среднее значение шума может изменяться в пределах «скользящего» окна. В данной работе предположим, что среднее значение шума в пределах каждого «скользящего» окна может меняться либо по линейному закону, либо по квадратичному закону:

$$y_{k_i} = b_k \cdot i + c_k \quad (1)$$

$$y_{k_i} = a_k \cdot i^2 + b_k \cdot i + c_k \quad (2)$$

где a_k, b_k, c_k – параметры, описывающие изменение среднего значения шума в пределах k-го блока стабилизации уровня ложной тревоги; $i = 0,1...M-1$. Коэффициенты a_k, b_k, c_k вычисляются на основе метода наименьших квадратов [5].

Для выбора модели аппроксимации среднего значения шума применяется критерий минимума дисперсии [6], суть которого заключается в сравнении значений дисперсий линейной и квадратичной модели при аппроксимации среднего значения шума.

На рисунке 3 представлена структурная схема предлагаемого алгоритма СУЛТ с выбором модели аппроксимации среднего значения шума.

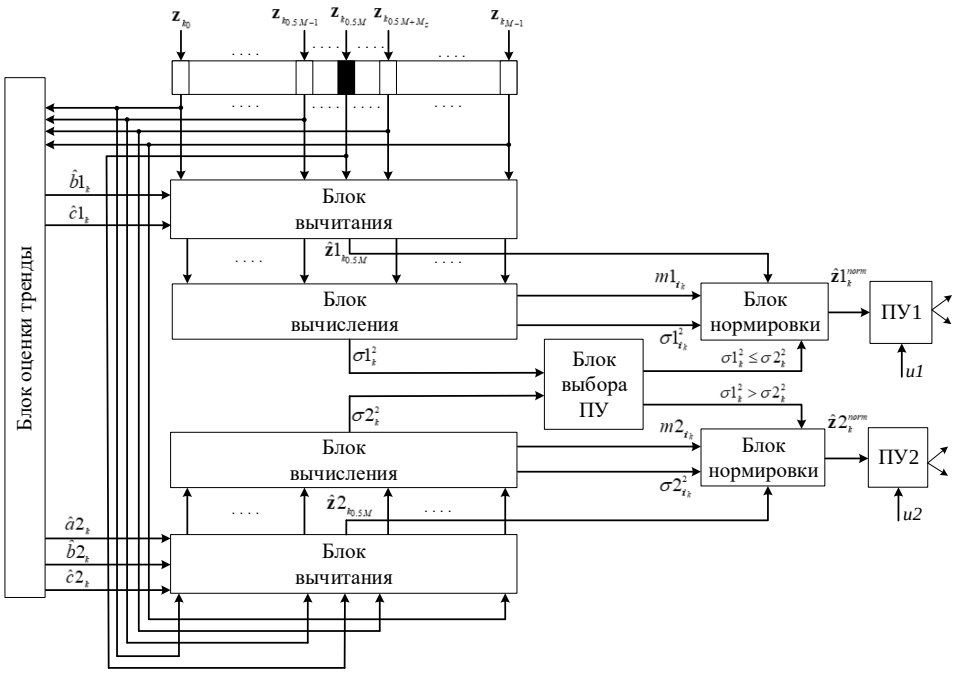


Рисунок 3 – Структурная схема предлагаемого алгоритма СУЛТ с выбором модели аппроксимации среднего значения шума

На рисунке 3 введены следующие обозначения: ПУ1 и ПУ2 – пороговые устройства; *u1* и *u2* – пороги обнаружения для каждой предлагаемой модели аппроксимации среднего значения шума; M_z — величина защитного интервала. Блоки на схеме выполняют следующие функции:

Блок оценки тренда: оценка коэффициентов $\hat{b}1_k$, $\hat{c}1_k$, $\hat{a}2_k$, $\hat{b}2_k$, $\hat{c}2_k$ с помощью метода наименьших квадратов; Блок вычитания: вычитание значения между элементом массива $\hat{\mathbf{z}}_k$ и оценкой линейного и квадратичного трендов (1) и (2); Блок вычисления: оценка математического ожидания и дисперсии входных отсчётов, за исключением отсчётов, входящих в «защитный» интервал; Блок нормировки: нормировка случайной величины на входе порогового устройства; Блок выбора ПУ: выбор порогового устройства ПУ1, либо ПУ2;

Предлагаемый алгоритм описывается следующими шагами:

а) оценка коэффициентов $\hat{b}1_k$, $\hat{c}1_k$, $\hat{a}2_k$, $\hat{b}2_k$, $\hat{c}2_k$ с помощью метода наименьших квадратов.

б) определение элементов массива $\hat{\mathbf{z}}1_k$ и $\hat{\mathbf{z}}2_k$ в «блоках вычитания»:

$$\hat{\mathbf{z}}1_{k_i} = \mathbf{z}_{k_i} - (\hat{b}1_k \cdot i + \hat{c}1_k), \tag{3}$$

$$\hat{\mathbf{z}}2_{k_i} = \mathbf{z}_{k_i} - (\hat{a}2_k \cdot i^2 + \hat{b}2_k \cdot i + \hat{c}2_k). \tag{4}$$

в) вычисление дисперсии $\sigma 1_k^2$ и $\sigma 2_k^2$ в «Блоках вычисления»

$$\sigma 1_k^2 = \sum_{i=1}^{M} \left(\mathbf{z}_{k_i} - (\hat{b}1_k \cdot i + \hat{c}1_k) \right)^2, \tag{5}$$

$$\sigma 2_k^2 = \sum_{i=1}^{M} \left(\mathbf{z}_{k_i} - (\hat{a}2_k \cdot i^2 + \hat{b}2_k \cdot i + \hat{c}2_k) \right)^2. \tag{6}$$

г) сравнение значений дисперсий $\sigma 1_k^2$ и $\sigma 2_k^2$.

Сравнение значений дисперсий $\sigma 1_k^2$ и $\sigma 2_k^2$ выполняется, если $\sigma 1_k^2 \le \sigma 2_k^2$ выбирается ПУ1 для обнаружения, а при $\sigma 1_k^2 > \sigma 2_k^2$ выбирается пороговое устройство ПУ2.

д) Определение нормированных случайных величин $\hat{\mathbf{z}}1_k^{norm}$ и $\hat{\mathbf{z}}2_k^{norm}$ на входе пороговых устройств ПУ1 и ПУ2

Нормированные случайные величины $\hat{\mathbf{z}}1_k^{norm}$ и $\hat{\mathbf{z}}2_k^{norm}$ определяются по выражениям:

$$\hat{\mathbf{z}}1_k^{norm} = \frac{\hat{\mathbf{z}}1_{k_{0.5M}} - m1_{\hat{\mathbf{z}}_k}}{\sqrt{\hat{\sigma}1^2{}_{\mathbf{z}_k}}}, \tag{7}$$

$$\hat{\mathbf{z}}2_k^{norm} = \frac{\hat{\mathbf{z}}2_{k_{0.5M}} - m2_{\hat{\mathbf{z}}_k}}{\sqrt{\hat{\sigma}2^2{}_{\mathbf{z}_k}}}, \tag{8}$$

где $m1_{\hat{\mathbf{z}}_k}, m2_{\hat{\mathbf{z}}_k}$ - математическое ожидание отсчётов на выходе «Блока вычитания»; $\hat{\sigma}1^2{}_{\mathbf{z}_k}$, $\hat{\sigma}2^2{}_{\mathbf{z}_k}$ - дисперсия отсчётов на выходе «Блока вычитания».

Величины $m1_{\hat{\mathbf{z}}_k}$, $m2_{\hat{\mathbf{z}}_k}$ определяются согласно выражениям:

$$m1_{\hat{\mathbf{z}}_k} = \frac{1}{M - (2M_z + 1)} \cdot \left(\sum_{i=0}^{0.5M - M_z} \hat{\mathbf{z}}1_{k_i} + \sum_{i=0.5M + M_z}^{M-1} \hat{\mathbf{z}}1_{k_i} \right) \tag{9}$$

$$m2_{\hat{\mathbf{z}}_k} = \frac{1}{M - (2M_z + 1)} \cdot \left(\sum_{i=0}^{0.5M - M_z} \hat{\mathbf{z}}2_{k_i} + \sum_{i=0.5M + M_z}^{M-1} \hat{\mathbf{z}}2_{k_i} \right). \tag{10}$$

е) Определение порогов *u*1 и *u*2 для обнаружения

В данной работе используем ряд Лагерра [7] для аппроксимации плотности распределения вероятностей величин $\hat{\mathbf{z}}1_k^{norm}$ и $\hat{\mathbf{z}}2_k^{norm}$. Пороги *u*1 и *u*2 обнаружения определятся по следующими выражениями [7, 8]:

$$F = \int_{u1}^{\infty} p_{\hat{\mathbf{z}}1_k^{norm}}(y)dy = \int_{u1}^{\infty} \frac{(y-q)^{v} e^{-(y-q)/d}}{d^{v+1}} \times$$
$$\times \sum_{j=0}^{m} \left[\sum_{k=0}^{j} \frac{(-1)^k j! \mu_{x,j-k}}{\Gamma(1+j-k+v)(j-k)!k!} \sum_{k=0}^{j} \frac{(-1)^k \Gamma(1+j+v) x^{j-k}}{\Gamma(1+j-k+v)(j-k)!k!} \right] dy, \quad (11)$$

$$F = \int_{u2}^{\infty} p_{\hat{\mathbf{z}}2_k^{norm}}(y)dy = \int_{u2}^{\infty} \frac{(y-q)^{v} e^{-(y-q)/d}}{d^{v+1}} \times$$
$$\times \sum_{j=0}^{m} \left[\sum_{k=0}^{j} \frac{(-1)^k j! \mu_{x,j-k}}{\Gamma(1+j-k+v)(j-k)!k!} \sum_{k=0}^{j} \frac{(-1)^k \Gamma(1+j+v) x^{j-k}}{\Gamma(1+j-k+v)(j-k)!k!} \right] dy, \quad (12)$$

где, F – вероятность ложной тревоги на канал обнаружения; $\Gamma(\bullet)$ – гамма-функция; q – минимальное значение массива случайной величины; $d = \frac{-\mu_{y,1}^{\ 2} + \mu_{y,2}}{-q + \mu_{y,1}}$, где $\mu_{y,t}$ – t-й момент случайной величины y; $v = \frac{\mu_{y,1} - q}{d} - 1$; $\delta_j = \sum_{k=0}^{j} \frac{(-1)^k j! \mu_{x,j-k}}{\Gamma(1+j-k+v)(j-k)!k!}$, где $\mu_{x,t}$ это t-й момент случайной величины $x = \frac{y-q}{d}$.

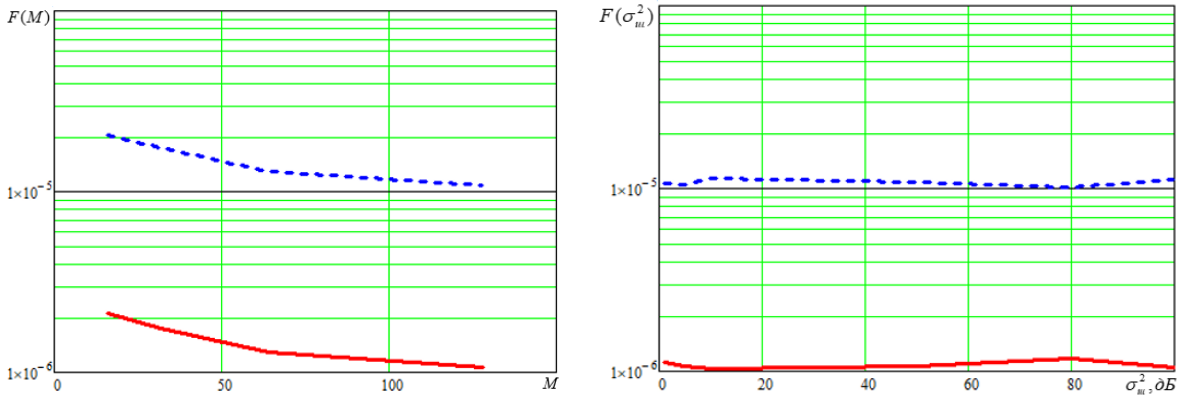
Обнаружение объекта в k-м «скользящем» окне выполняется в условиях: $\hat{\mathbf{z}}1_k^{norm} > u1$ если $\sigma 1_k^2 \le \sigma 2_k^2$ или $\hat{\mathbf{z}}2_k^{norm} > u2$ если $\sigma 1_k^2 > \sigma 2_k^2$.

Результаты моделирования для предлагаемого алгоритма

Эффективность предлагаемого алгоритма определяется с помощью имитационного моделирования при следующих параметрах:

- максимальное число каналов обнаружения: $N_{\max} = 256$;
- вероятность ложной тревоги на канал обнаружения $F_{зад}$=10^{-5}, 10^{-6}, 10^{-7};
- закон распределение шума в каналах обнаружения: экспоненциальный.

Зависимость вероятности ложной тревоги F от размера «скользящего» окна M и изменения дисперсии шума $\sigma_ш^2$ показана на рисунке 4а и 4б соответственно.



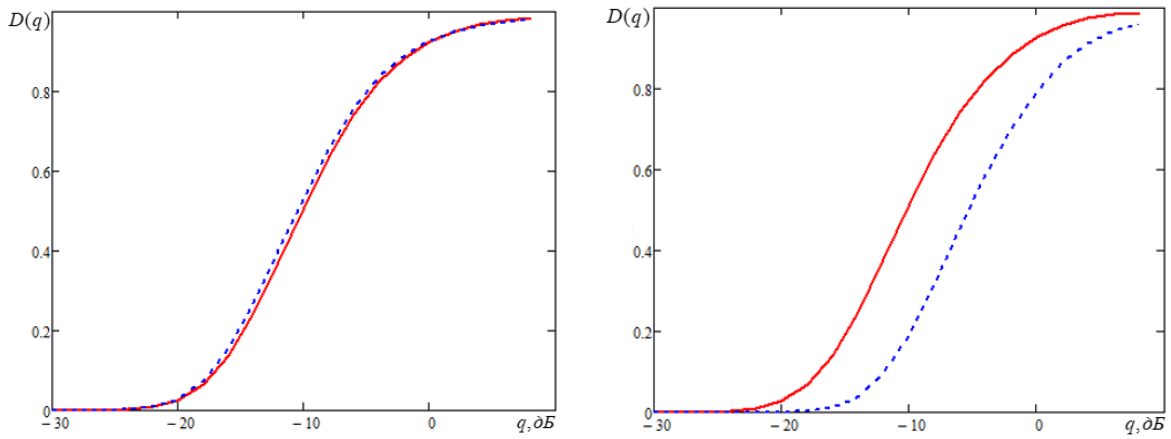
а б

Рисунок 4 – Зависимость вероятности ложной тревоги F от размера «скользящего» окна M и от изменения дисперсии шума $\sigma_ш^2$

Параметры, при которых построены данные зависимости: заданные вероятности ложной тревоги F=10^{-5}, F=10^{-6}; объём выборки меняется в диапазоне M = 16…128. Сплошной линией показана зависимость при значении F=10^{-6}; пунктирной линией – значение F=10^{-5}. Анализ зависимостей на рисунке 4а показывает, что вероятность ложной тревоги стремится к

заданному значению $F_{зад}$ с увеличением объема выборки M. Анализ зависимостей на рисунке 4б показывает, что вероятность ложной тревоги F постоянна при изменении значений дисперсии шума $\sigma_{ш}^2$.

На рисунке 5 показаны характеристики обнаружения предлагаемого алгоритма СУЛТ при разных значениях a, b, c. Вероятность ложной тревоги равны $F=10^{-7}$ и объём выборки M =128.



а б

Рисунок 5 – Характеристики обнаружения предлагаемого алгоритма СУЛТ при разных значениях a, b и c:
а) a=0; b=0; c=0;
б) a=0,1; b=2; c=3

На рисунках 5 сплошная линия соответствует характеристикам обнаружения предлагаемого алгоритма СУЛТ; пунктирная линия – алгоритму УС-СУЛТ.

Анализ зависимостей, представленных на рисунках 5а, показывает, что для предлагаемого алгоритма при вероятности правильного обнаружения D=0,9 в случае F=10^{-7} проигрыш в пороговом отношении сигнал-шум по сравнению с алгоритмом УС-СУЛТ примерно составляет 0,15 *дБ*. Для зависимостей на рисунке 5б выигрыш в пороговом отношении сигнал-шум предлагаемого алгоритма по сравнению с алгоритмом УС-СУЛТ примерно составляет 3,65 *дБ*.

В таблице 1 приведён выигрыш в пороговом отношении сигнал-шум предлагаемого алгоритма по сравнению с алгоритмом УС-СУЛТ для различных параметров a, b и c в случае F=10^{-7}.

Таблица 1. Выигрыш в пороговом отношении сигнал-шум, *дБ*

Модели аппроксимации среднего значения шума	Параметры			Заданная вероятность ложной тревоги, F=10^{-7}
	a	b	c	
Линейный	0	1,5	1,5	1,83
	0	2,0	1,0	1,54
	0	3,0	5,0	3,52
Квадратный	0,15	1,5	2,0	4,25
	0,25	3,0	4,5	5,12
	0,3	5,0	3,0	5,78

Анализ данных таблицы 1 показывает, что для двух моделей аппроксимации среднего значения шума при изменении значений a, b и c выигрыш в пороговом отношении сигнал-шум предлагаемого алгоритма изменяется и составляет 5,78 *дБ* при значении a=0,3 и b=5,0 и c=3,0.

Заключение

Таким образом, в работе рассмотрен алгоритм стабилизации уровня ложной тревоги с выбором модели аппроксимации среднего значения шума в пределах «скользящего» окна. В основе предлагаемого алгоритма лежит выбор линейной или квадратичной аппроксимации на основе сравнения оценки дисперсий. Результаты моделирования показывают, что предлагаемый алгоритм обеспечивает выигрыш в пороговом отношении сигнал-шум до 5,78 *дБ* по сравнению с известным алгоритмом УС-СУЛТ. Вероятность ложной тревоги стремится к заданному значению $F_{зад}$ с увеличением размера «скользящего» окна и постоянна при изменении значений дисперсии шума $\sigma_{ш}^2$.

Литература

1. O.A. Stepanov. Fundamentals of estimation theory with applications to navigation information processing problems // SPb.: State Scientific Center of the Russian Federation JSC Concern Central Research Institute Elektropribor. 2010. 509 p.

2. П.А. Бакулев. Радиолокационные системы // Учебник для ВУЗов. М.: Радиотехника. 2004. 325 с.

3. S. Himonas. Optimum detection of rayleigh signals in nonstationary noise // Tenth Annual International Phoenix Conference on Computers and Communications. 1991. P. 401-405.

4. M.A. Khalighi and M.H. Bastani. Adaptive CFAR processor for nonhomogeneous environments // IEEE Transactions on Aerospace and Electronic Systems. vol. 36. №. 3. pp. 889-897. July 2000.

5. H.M. Finn, R.S. Johnson. Adaptive detection model with threshold control as a function of spatially sampled clutter-level estimates // RCA.1968. pp. 414–464.

6. Дж. Прокис. Цифровая связь. Пер. с англ // М.: Радио и связь. 2000. 800 с.

7. G. Alexits. Convergence Problems of Orthogonal Series // Oxford: Pergamon. 1961. 350 с.

8. В.Т. Горяинов, А.Г. Журавлёв, В.И. Тихонов. Статистическая радиотехника: примеры и задачи // М.: Советское радио. 1980. 450 с.

STUDY OF THE PERFORMANCE OF TE ALGORITHM FOR FALSE ALARM LEVER STABILISATION WITH THE CHOICE MODEL OF APPROXIMATION OF THE MEAN NOISE VALUE

Dr. Sc., Prof. Belokurov V.A.[1], Asp. Nguyen T.Q.[1]

[1]Ryazan State Radio Engineering University named after V.F. Utkin

This paper analyzes the performance the algorithm for false alarm lever stabilisation with the choice model of approximation of the mean noise value. The performance of the proposed algorithm against the background of non-stationary noise in comparison with the CA-CFAR algorithm. The analysis results showed that the proposed algorithm provides a gain in the threshold signal-to-noise ratio up to 5,78 dB compared to the CA-CFAR algorithm.

ИЗМЕРЕНИЕ РАССТОЯНИЯ ЧМ ДАЛЬНОМЕРОМ ПРИ НАЛИЧИИ МЕШАЮЩИХ ОТРАЖЕНИЙ

проф. Паршин В.С., асп. Нгуен В.Д.

Рязанский государственный радиотехнический университет им. В.Ф. Уткина

На практике, в системах контроля над состоянием технологического оборудования широкое применение нашли дальномеры с частотной модуляцией излучаемого сигнала. Они обеспечивают высокую точность измерения расстояния до отражающей поверхности (менее мм в беспомеховой обстановке), малочувствительны к загазованности и запыленности среды распространения радиоволн. Существенным фактором, который ограничивает применение таких дальномеров, является наличие мешающих отражений. В работе проведен анализ алгоритмов, позволяющих уменьшить влияние помех различной интенсивности (в том числе при отношении сигнал-помеха меньше единицы), основанных на варьировании начальной фазы сигнала разностной частоты и на особенностях его функции правдоподобия. Приведены результаты обработки экспериментальных данных, полученных с помощью серийно выпускаемых дальномеров.

Измерение расстояния до некоторой отражающей поверхности при использовании дальномеров с частотной модуляцией излучаемого сигнала (ЧМ дальномер) часто сопровождается наличием мешающих отражений (МО). Причина появления МО - на пути распространения электромагнитных волн могут находиться различные отражающие элементы. Ограничиваясь случаем наличия одного МО, сигнал разностной частоты (СРЧ), снимаемый на выходе смесителя ЧМ дальномера, на интервале $(0,T_{мод}/2)$ можно представить в виде [1]:

$$y(t)=S_c\cos\left[\omega_0\frac{2R}{c}+\frac{4\Delta\omega_d R}{cT_{мод}}t+\varphi_c\right]+S_n\cos\left[\omega_0\frac{2R_n}{c}+\frac{4\Delta\omega_d R_n}{cT_{мод}}t+\varphi_n\right]+\zeta(t), \qquad (1)$$

где S_c, S_n - соответственно амплитуды полезного сигнала и сигнала, отраженного от помехи; ω_0 - несущая частота генератора; $\Delta\omega_d$ - диапазон перестройки несущей частоты; $T_{мод}$ - период модуляции несущего колебания; R, R_n - расстояния до отражающей поверхности и помехи; φ_c, φ_n - фазы сигналов, соответствующие полезному сигналу и МО; c - скорость света; $\zeta(t)$ - белый нормальный шум со спектральной плотностью N_0.

Как правило, в современных ЧМ дальномерах за оценку $\hat{\omega}$ частоты СРЧ принимается та частота, на которой находится максимальная спектральная составляющая спектральной плотности амплитуды $S(\omega)$ СРЧ, то есть [2,3]

$$S(\hat{\omega})=\max_{\omega}\{S(\omega)\}. \qquad (2)$$

Оценка расстояния определяется как $\hat{R}=\hat{\omega}T_{мод}c/4\Delta\omega_d$.Использование алгоритма (2) при наличии МО приводит к увеличению погрешности измерения. Например, при отношении сигнал-помеха 2 дБ и диапазоне перестройки 500 МГц погрешность измерения примерно равна 18 см.

Ниже представлены результаты обработки экспериментальных данных, полученные с помощью уровнемеров БАРС 351, 352, выпускаемых в городе Рязань на ООО предприятие «Контакт1». Измерения проведены на сертифицированном измерительном стенде [4]. Диапазон перестройки несущей равен 500 МГц. Несущая частота равна 10 ГГц. При проведении эксперимента радиолокационные отражатели, имитирующие отражающую поверхность и МО, перемещались с шагом 3 мм.

1. Измерение расстояния с помощью варьирования начальной фазы СРЧ

Известно, что график погрешности оценки расстояния, возникающей под воздействием МО, имеет колебательный характер [4] и периодически переходит через точки с нулевой погрешностью тогда, когда взаимная энергия полезного сигнала и МО равна нулю. Уменьшить влияние МО

можно за счет варьирования несущей частоты [4,5] при сохранении диапазона ее перестройки. Согласно (1) при таком варьировании будут изменяться начальные фазы СРЧ, соответствующих отражениям от отражающей поверхности и МО. Вследствие этого будет изменяться энергия СРЧ. Оптимальная несущая частота, при которой влияние МО уменьшается, равна [4,5] полусумме значений тех несущих частот, при которых полная энергия СРЧ минимальна и максимальна. Недостатком такого способа являются повышенные требования при реализации передатчика ЧМ дальномера.

В работе предлагается изменять начальную фазу СРЧ без варьирования несущей частоты передатчика. Для этого СРЧ $y(k)$ длиной N отсчетов разбивается на два блока длиной M и $(N-M)$ отсчетов каждый. Вычисляется энергия $E_{СРЧ}$ СРЧ с использованием первого блока длиной M отсчетов. Затем СРЧ сдвигается «вправо» на один отсчет. Первым отсчетом блока длиной M отсчетов будет его второй отчет, а последним – первый отсчет второго блока длиной $(N-M)$. После вычисления энергии при сдвиге на один отсчет первый блок снова сдвигается еще на один отсчет «вправо». Недостающий отсчет берется из второго блока и снова вычисляется энергия СРЧ. В результате вычислений получаем

$$E_{СРЧ}(m)=\sum_{k=1+m}^{M+m}\left[y^2(k)\right]=\sum_{k=1+m}^{M+m}\left[s_c^2(k)+s_n^2(k)+2s_c(k)s_n(k)\right]=E_c+E_n+E_{cn},m=0,1,2,\ldots, \quad (3)$$

где $s_c(k)$, $s_n(k)$ - соответственно полезный сигнал и сигнал, отраженный от МО; E_c - энергия полезного сигнала; E_n - энергия сигнала, отраженного от помехи; E_{cn} - взаимная энергия полезного сигнала и сигнала, отраженного от помехи.

Вычисления со сдвигом «вправо» на один отсчет производятся до тех пор, пока не будут определены максимальное i_1 и минимальное i_2 значения энергии СРЧ. Поскольку при сдвиге первого блока меняются начальные фазы полезного сигнала и помехи, то функция $E_{СРЧ}(m)$ имеет колебательный характер. При сдвиге отсчетов первого блока СРЧ на величину

$$i_{опт}=(i_1+i_2)/2. \quad (4)$$

Взаимная энергия будет минимальна, то есть помеха на полезный сигнал будет оказывать существенно меньшее влияние.

Указанные выше операции должны производиться при каждом измерении, поскольку взаимная энергия СРЧ зависит от расстояния. Необходимо отметить, что если исходный СРЧ разбивается на два блока, то это эквивалентно уменьшению диапазона перестройки несущей частоты пропорционально уменьшению длительности сигнала.

На рисунках 1, 2 показаны типичные мгновенные погрешности измерения расстояния, вычисленные так:

$$\Delta R=\hat{R}-R_{ист}, \quad (5)$$

где $\hat{R}$ - результат измерения расстояния; $R_{ист}$ - истинное расстояние.

На рисунке 1 показана погрешность оценки расстояния при наличии МО, который находится на расстоянии 7,7 метра. График 1 получен с помощью алгоритма (2). График 2 получен при варьировании начальной фазы с последующим использованием алгоритма (2). Общее число отсчетов СРЧ ($N=1024$) при вычислениях разбивалось на два блока по 512 отсчетов. Отношение сигнал-помеха равно 4 дБ.

На рисунке 2 показана погрешность измерения малых расстояний. МО здесь является спектральная плотность СРЧ, вычисленная на отрицательных частотах – своеобразная помеха, от которой невозможно избавиться. График 1 получен с помощью алгоритма (2). График 2 получен при той начальной фазе, при которой взаимная энергия спектральных плотностей на

положительных и отрицательных частотах минимальна. Из рисунка 2 следует, что погрешность измерения существенно уменьшается - примерно в 2-3 раза. Затраты на реализацию алгоритма с поиском минимума взаимной энергии сигнала и МО минимальны. Эта операция выполняется программно. При измерении больших расстояний, как видно из рисунка 1, уменьшение погрешности невелико. Это можно объяснить влиянием паразитной амплитудной модуляции СРЧ и его нелинейными искажениями.

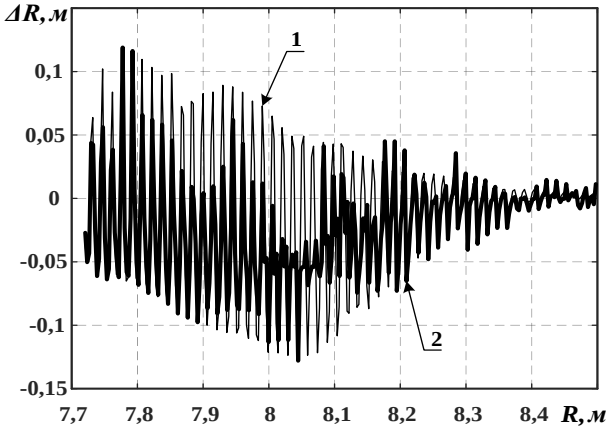


Рисунок 1 - Мгновенные погрешности при измерении больших расстояний

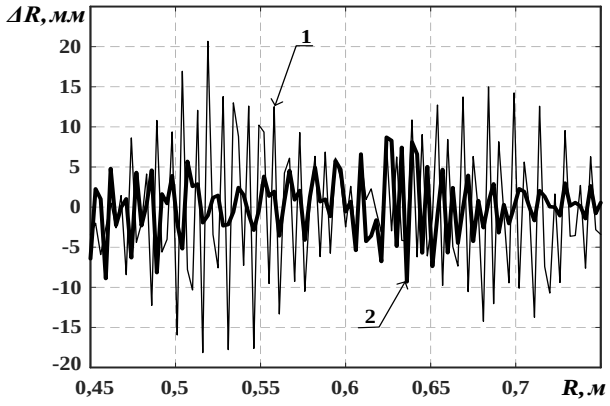


Рисунок 2 - Мгновенные погрешности при измерении малых расстояний

2. Измерение расстояния при отношении сигнал-помеха, меньшем единицы

Такая ситуация встречается, например, при измерении расстояния до уровня радиопрозрачной жидкости, находящейся вблизи дна резервуара. В этом случае использование алгоритма (2) приводит к слишком большим погрешностям, поскольку интенсивность спектральной плотности СРЧ, отраженного от дна, будет превышать интенсивность спектральной плотности полезного сигнала. Для уменьшения ошибки измерения в [6] предложена комбинация компенсации помехи и следящего алгоритма, работа которого подробно изложена [4].

Логарифм функции отношения правдоподобия (ЛФОП) СРЧ на интервале $(0,T_{мод}/2)$ можно представить в виде [2,3]

$$ln\left\{\Lambda\left[S_c,R,\varphi_c,S_n,R_n,\varphi_n\right]\right\}=\frac{2}{N_0}\int_0^{T_{мод}/2}\left\{y(t)S_{on}\left[t,S_{on},\varphi_{on},R_{on}\right]-\frac{1}{2}S_{on}^2\left[t,S_{on},\varphi_{on},R_{on}\right]\right\}dt, \qquad (6)$$

где $S_{on}\left[t,S_{on},\varphi_{on},R_{on}\right]$ - опорный сигнал, структура которого определяется моделью (1).

Вычисляя соответствующий интеграл в (6), пренебрегая слагаемыми с удвоенной частотой, получим, что сигнальная функция (СФ) будет равна сумме СФ полезного сигнала и СФ помехи, то есть

$$\begin{aligned}&q\left(S_c,R,\varphi_c,S_n,R_n,\varphi_n\right)=q_c\left(S_c,R,\varphi_c\right)+q_n\left(S_n,R_n,\varphi_n\right)=\\&\frac{S_cS_{on}T_{мод}}{2N_0}cos\left[\omega_0\left(\frac{2R}{c}-\frac{2R_{on}}{c}\right)+\varphi_c\left(\frac{2R}{c}\right)-\varphi_{on}\left(\frac{2R_{on}}{c}\right)\right]\times\frac{sin\left\{\Delta\omega_d\left(R-R_{on}\right)/c\right\}}{\Delta\omega_d\left(R-R_{on}\right)/c}+\\&+\frac{S_nS_{on}T_{мод}}{2N_0}cos\left[\omega_0\left(\frac{2R_n}{c}-\frac{2R_{on}}{c}\right)+\varphi_n\left(\frac{2R_n}{c}\right)-\varphi_{on}\left(\frac{2R_{on}}{c}\right)\right]\times\frac{sin\left\{\Delta\omega_d\left(R_n-R_{on}\right)/c\right\}}{\Delta\omega_d\left(R_n-R_{on}\right)/c}.\end{aligned} \qquad (7)$$

Функция (7) является многоэкстремальной. В том случае, когда расстояние между помехой и полезным сигналом достаточно большое, СФ полезного сигнала и СФ помехи практически не оказывают влияния друг на друга. Наибольшие экстремумы СФ полезного сигнала и СФ помехи

будут соответствовать расстояниям до полезного отражателя и МО. При малых расстояниях между МО и полезным отражателем из-за взаимного влияния СФ их максимальные экстремумы не будут соответствовать соответствующим расстояниям. В [4,6] предложен и проведен анализ следящего алгоритма за экстремумами функции (7). Слежение за экстремумом ЛФОП (6), который соответствует МО, позволяет оценить его параметры - амплитуду и местоположение. Амплитуда сигнала оценивалась методом максимального правдоподобия [2,3]. За оценку его местоположения принималось та величина, при которой величина соответствующего экстремума была наибольшей. Фазу φ_n будем считать известной. В результате возможно сформировать компенсирующий сигнал $S_к(t)$, параметрами которого являются полученные оценки $\hat{S}_n$, $\hat{R}_n$. Результирующий СРЧ запишется так:

$$S_p(t) = y(t) - S_к(t) = S_c cos\left[\omega_0 2R/c + 4\Delta\omega_d Rt/cT_{мод} + \varphi_c\right] + \\ + S_n cos\left[\omega_0 2R_n/c + 4\Delta\omega_d R_n t/cT_{мод} + \varphi_n\right] - \hat{S}_n cos\left[\omega_0 2\hat{R}_n/c + 4\Delta\omega_d \hat{R}_n t/cT_{мод} + \varphi_n\right] + \zeta(t). \quad (8)$$

Слежение за экстремумом ЛФОП (6), который соответствует расстоянию до пленного отражателя, позволяет получить оценку расстояния до полезного отражателя. Разумеется, полной компенсации помехи не произойдет из-за паразитной амплитудной модуляции и нелинейных искажений СРЧ, различного рода переотражений, погрешностей при оценке параметров помехи. Главное при компенсации - обеспечить [4] отношение сигнал-помеха более 2 дБ.

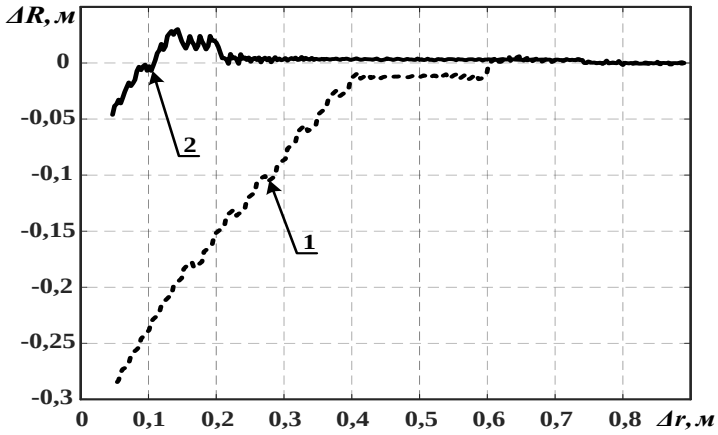


Рисунок 3 - Мгновенные погрешности измерения расстояния до отражающей поверхности вблизи дна

На рисунке 3 приведены результаты обработки экспериментальных данных. Отношение сигнал-помеха составляло величину -6 дБ. Под Δr понимается расстояние между МО и полезным отражателем. МО располагается на расстоянии 10 метров от ЧМ дальномера.

График 1 получен при оценке расстояния при использовании следящего алгоритма без компенсации. График 2 получен при тех же условиях, но с компенсацией помехи. Полезный отражатель перемещался с шагом 2 мм, а МО - с шагом 1 мм в противоположном от полезного отражателя направлении. Из рисунка 3 видно, что при компенсации помехи погрешность измерения существенно уменьшается. Срыв слежения наступает на существенно меньших расстояниях между МО и полезным сигналом. Проведено сравнение полученных результатов с результатами измерения расстояния при наличии МО с использованием методов параметрического спектрального анализа [4,7]. Проведенное сравнение позволяет утверждать, что по крайней мере алгоритм, основанный на использовании ЛФОП в комбинации с компенсацией помехи не уступает последним при значительно более простой аппаратурной реализации. Ограничение при его использовании – величина перемещения отражающей поверхности при двух последовательных измерениях не должна [4] превышать величины четверти длины волны несущего колебания ($\lambda/4$).

3. Оценка фазовой характеристики ЧМ дальномера

Использование следящего алгоритма, реализующего ЛФОП, позволяет существенно уменьшить погрешность измерения. Однако для его применения требуется знать фазу φ_c, зависящую от расстояния, то есть необходимо знать фазовую характеристику (ФХ) дальномера. В [4] подробно изложена методика определения ФХ. Согласно [4] значение ФХ для каждого расстояния определяется так:

$$\varphi_c = \hat{\varphi}_н - \omega_0 2\hat{R}/c\,, \qquad (9)$$

где $\hat{\varphi}_н = -arctg\left\{ImS(i\hat{\omega})/ReS(i\hat{\omega})\right\}$ - начальная фаза СРЧ.

Из (9) следует, что для определения ФХ необходимо предварительно оценивать расстояние до отражающей поверхности. Использование для этого алгоритма (2) приводит к слишком большим ошибкам. В [8] предложено для оценки ФХ использовать ЛФОП [8]. Его «узкие» экстремумы [4] позволяют получать оценку фазу с меньшими погрешностям, поскольку погрешность измерения расстояния с его использованием будет меньше. Ниже изложена методика оценки ФХ на сертифицированном измерительном стенде [4], что можно совместить с калибровкой ЧМ дальномера. Последовательность действия на практике следующая:

1. При каждом измерении необходимо определить предварительную оценку расстояния $\hat{R}$, используя алгоритм (2), гарантирующий попадание оценки в окрестность глобального максимума ЛФОП. Погрешность измерения расстояния на первом этапе не должна превышать $\lambda/4$, т.е. должно выполняться очевидное условие [4]

$$-\lambda/4 \le \hat{R} - R_{ист} \le \lambda/4\,. \qquad (10)$$

2. Для каждого измерения расстояния, рассчитанного на первом этапе по алгоритму (2), поиск ФХ СРЧ осуществляется в интервале расстояний, удовлетворяющем условию (10). Варьированием расстояния и фазы опорного сигнала необходимо добиться максимизации экстремума, который попадает в интервал (10). За оценку величины ФХ принимается то ее значение, при котором выполняется условие

$$\int_{\hat{R}-\lambda/4}^{\hat{R}} \Lambda(R,\varphi_c)\partial R = \int_{\hat{R}}^{\hat{R}+\lambda/4} \Lambda(R,\varphi_c)\partial R\,, \qquad (11)$$

где $\Lambda(R,\varphi_c)$ - ЛФОП.

Результаты обработки экспериментальных данных показаны на рисунке 4. Эксперимент проведен при следующих условиях. РЛО перемещался на специальной тележке вдоль измерительной ленты от 0,85м до 8,3м с шагом 0,05м. Из конструктивных особенностей элементов измерительного стенда, отражений от стен и потолка стендового зала возникают различные переотражения – множественные МО малой интенсивности (отношение сигнал – помеха больше 40 дБ). Избавиться от таких МО на практике чрезвычайно сложно. График 1 получен на основе соотношения (9), а график 2 в соответствии с (11).

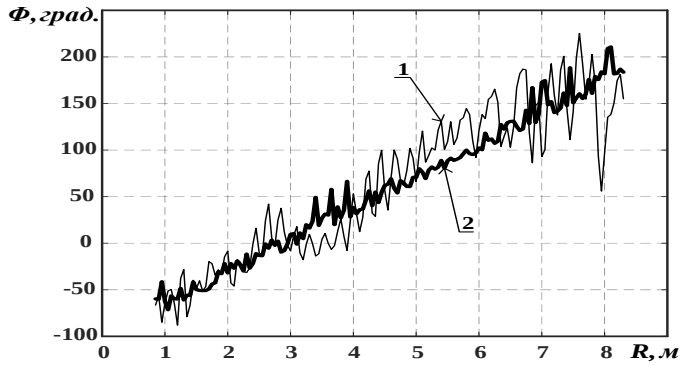


Рисунок 4 - Оценка фазовой характеристики дальномера

Погрешность оценки ФХ обусловлена двумя причинами. Это влияние МО малой интенсивности и неточность установки радиолокационного отражателя вдоль измерительной ленты (ее цена деления 1 мм).

Из анализа рисунка 4 следует, что оценка ФХ с использованием (11) позволяет уменьшить её погрешность, что позволяет снизить смещения при использовании алгоритма (11).

Литература

1. Виницкий А.С. Очерк основ радиолокации при непрерывном излучении радиоволн. - М.: Советское радио, 1961. - 495 с.

2. Тихонов В.И. Оптимальный прием сигналов. – М.: Радио и связь, 1983, – 320 с.

3. Куликов Е.И., Трифонов А.П. Оценка параметров сигналов на фоне помех. – М.: Сов. радио, 1978, – 296 с.

4. Атаянц Б.А., Езерский В.В., Давыдочкин В.М., Паршин В.С., Смольский С.М. Прецизионные системы ближней частотной радиолокации промышленного применения – М: Радиотехника, 2012. -512 с.

5. Паршин, В.С. Уменьшение погрешности измерения расстояния дальномером с частотной модуляцией зондирующего сигнала при использовании перестройки несущей частоты передатчика /В.С. Паршин, Н.С. Заигров // Вестник Рязанского государственного радиотехнического университета. -2020. № 74. – С. 14-22. DOI 10.21667/1995-4565-2020-74-14-22. – EDNIBJKEX.

6. Паршин, В. С. Измерение расстояния до отражающей поверхности ЧМ дальномером при отношении сигнал/помеха меньше единицы / В. С. Паршин, В. Д. Нгуен // Цифровая обработка сигналов и ее применение DSPA - 2024 (Серия: Цифровая обработка сигналов и её применение): Доклады XXVI Международной конференции, Москва, 27–29 марта 2024 года. – Москва: Российское научно-техническое общество радиотехники, электроники и связи им. А.С. Попова, 2024. – С. 113-118. – EDN BEXCFO.

7. Марпл С.Л. Цифровой спектральный анализ и его приложения: Пер. с англ./ Под ред. И.С. Рыжика. – М.: Мир, 1990. – С. 584.

8. Паршин, В. С. Алгоритмы оценки фазовой характеристики дальномера с частотной модуляцией зондирующего сигнала / В. С. Паршин, В. Д. Нгуен // Цифровая обработка сигналов. – 2023. – № 2. – С. 22-26. – EDN NWHKGI.

MEASURING DISTANCE WITH A FM RANGE FINDER IN THE PRESENCE OF INTERFERING REFLECTIONS

Prof. Parshin V.S., grad. stud Nguyen V.D.

Ryazan State Radio Engineering University named after V.F. Utkin

In practice, rangefinders with frequency modulation of the emitted signal are widely used in systems for monitoring the condition of process equipment. They ensure high accuracy of measuring the distance to the reflecting surface (less than mm in a noise-free environment), and are insensitive to gas and dust contamination of the radio wave propagation environment. A significant factor that limits the use of such rangefinders is the presence of interfering reflections. The paper analyzes algorithms that reduce the impact of interference of varying intensity (including when the signal-to-interference ratio is less than one), based on varying the initial phase of the difference frequency signal and on the features of its likelihood function. The results of processing experimental data obtained using commercially available rangefinders are presented.

НЕЙРОСЕТЕВАЯ ВИНЕРОВСКАЯ ФИЛЬТРАЦИЯ В ОБЛАСТИ ДИСКРЕТНОГО ВЕЙВЛЕТ-ПРЕОБРАЗОВАНИЯ ДЛЯ УЛУЧШЕНИЯ КАЧЕСТВА РАСПОЗНАВАНИЯ ЦИФРОВЫХ ИЗОБРАЖЕНИЙ

асп. Алимагадов К.А., проф., д.ф.-м.н. Умняшкин С.В.

Национальный исследовательский университет «МИЭТ»

Исследование посвящено улучшению подавления шумов в цифровых изображениях, искажённых аддитивным белым гауссовым шумом. Разработан алгоритм нейросетевой винеровской фильтрации в области дискретного вейвлет-преобразования. Качество обработки оценивалось на основе сравнения значений метрик пикового отношения сигнал/шум и индекса структурного сходства. Предложенный алгоритм подавления шумов обеспечивает больший рост метрик качества по сравнению с рассмотренными аналогами. Помимо этого, было рассмотрено обучение нейросетевого детектора SqueezeDet для решения задачи распознавания автомобильных номеров с применением аугментаций, включающих зашумление входных изображений с последующей их фильтрацией разработанным алгоритмом. Результаты экспериментов по распознаванию автомобильных номеров на тестовом наборе данных продемонстрировали наибольшую точность при использовании предложенного алгоритма подавления шумов для обработки входных изображений нейросетевого детектора.

Введение

Подавление шумов в цифровых изображениях способствует улучшению визуального качества фото- и видео данных, а также снижению доли ошибок распознавания при последующем применении алгоритмов интеллектуального анализа данных [1]. На сегодняшний день наиболее качественные результаты подавления шумов демонстрируют методы обработки на основе нейронных сетей [2-4]. Одним из направлений исследований в области нейросетевой фильтрации изображений является разработка алгоритмов, сочетающих преимущества машинного обучения и классических методов цифровой обработки, что позволяет снизить количество параметров моделей и повысить их интерпретируемость [3, 5].

Целью настоящей работы является улучшение качества подавления шумов в цифровых изображениях и изучение влияния такой обработки на точность работы алгоритмов распознавания.

Разработана нейронная сеть, выполняющая винеровскую фильтрацию изображений в области дискретного вейвлет-преобразования (ДВП) с учётом контурной информации. Предлагаемый алгоритм подавления шумов перед фильтрацией выполняет нечёткую классификацию коэффициентов ДВП-спектра на два класса в зависимости от локализации базисных функций (двумерных вейвлетов) на изображении: 1) «контурные», соответствующие контурам на изображении, 2) «фоновые», соответствующие фоновым областям. Такая классификация позволяет гибко подстраивать параметры фильтра Винера к локальным особенностям обрабатываемых изображений.

Модель шума

В рамках исследования рассматривается модель белого аддитивного не коррелированного с изображением шума, распределённого по нормальному закону с нулевым математическим ожиданием $\mu = 0$ и среднеквадратическим отклонением (СКО) σ.

Дискретные полутоновые изображения будем представлять в виде матриц, элементы которых определяют яркость пикселей. Искажение изображения в формализованном виде описывается выражением:

$$\mathbf{X} = \mathbf{Y} + \mathbf{N},$$

где $\mathbf{X} = (x_{m,n})$, $\mathbf{Y} = (y_{m,n})$ и $\mathbf{N} = (\eta_{m,n})$ – матрицы зашумлённого изображения, исходного изображения и шума соответственно.

Дискретное вейвлет-преобразование

Фильтрацию изображений будем выполнять в базисе биортогональных фильтров 9/7 [6], импульсные характеристики (ИХ) которых задают одномерные свёртки, применяющиеся последовательно вдоль строк, затем вдоль столбцов матрицы изображения.

В процессе обработки выполняется 3-х уровневое ДВП без прореживания откликов ВЧ и НЧ-фильтров, задающих преобразование [6]. При вычислении каждого i-го ($i \in \{1, 2, 3\}$) уровня

ДВП получаются четыре набора вейвлет-коэффициентов (саббэндов): один саббэнд, соответствующий низкочастотной составляющей ДВП-спектра $\mathbf{V}_i^A$, и три саббэнда $\mathbf{V}_i^H$, $\mathbf{V}_i^V$, $\mathbf{V}_i^D$, соответствующие горизонтальной, вертикальной и диагональной высокочастотным составляющим. Фильтрация применяется только к саббэндам $\mathbf{V}_i^H$, $\mathbf{V}_i^V$, $\mathbf{V}_i^D$, а саббэнд $\mathbf{V}_i^A$ используется для вычисления (i+1)-го уровня ДВП.

Алгоритм подавления шума

Входными данными алгоритма являются изображение **X** и заданная величина СКО σ шума **N**. Общая схема разработанного алгоритма подавления шума состоит из следующих шагов.

Шаг 1. Вычисление ДВП-спектра входного изображения: **V** = ДВП(**X**).

Шаг 2. Оценка дисперсии шума в саббэндах $(\mathbf{V}_i^H, \mathbf{V}_i^V, \mathbf{V}_i^D)_{i=1}^3$ по величине σ.

Шаг 3. Определение степени принадлежности вейвлет-коэффициентов саббэндов $(\mathbf{V}_i^H, \mathbf{V}_i^V, \mathbf{V}_i^D)_{i=1}^3$ к классам «контурные» и «фоновые».

Шаг 4. Нахождение оценок спектров мощности полезного сигнала $\hat{\mathbf{S}}_\mathbf{Y} = (\hat{s}_{k,l})$ и шума $\hat{\mathbf{S}}_\mathbf{N} = (\hat{p}_{k,l})$ для каждого класса («контурных»/«фоновых») вейвлет-коэффициентов.

Шаг 5. Вычисление коэффициентов $\mathbf{W} = (w_{k,l})$ фильтра Винера для каждого класса вейвлет-коэффициентов.

Шаг 6. Фильтрация ДВП-спектра, состоящая в покомпонентном перемножении коэффициентов фильтра Винера и соответствующих им коэффициентов ДВП-спектра: $\hat{\mathbf{U}} = (\hat{u}_{k,l} = w_{k,l} \cdot v_{k,l})$.

Шаг 7. Восстановление изображения путём выполнения обратного ДВП: $\hat{\mathbf{Y}}$ = ОДВП($\hat{\mathbf{U}}$).

Шаг 8. Компенсация артефактов, возникших в результате винеровской фильтрации.

Перечисленные выше шаги были реализованы как операции, выполняемые слоями нейронной сети с архитектурой, которая укрупнённо представлена на рисунке 1.

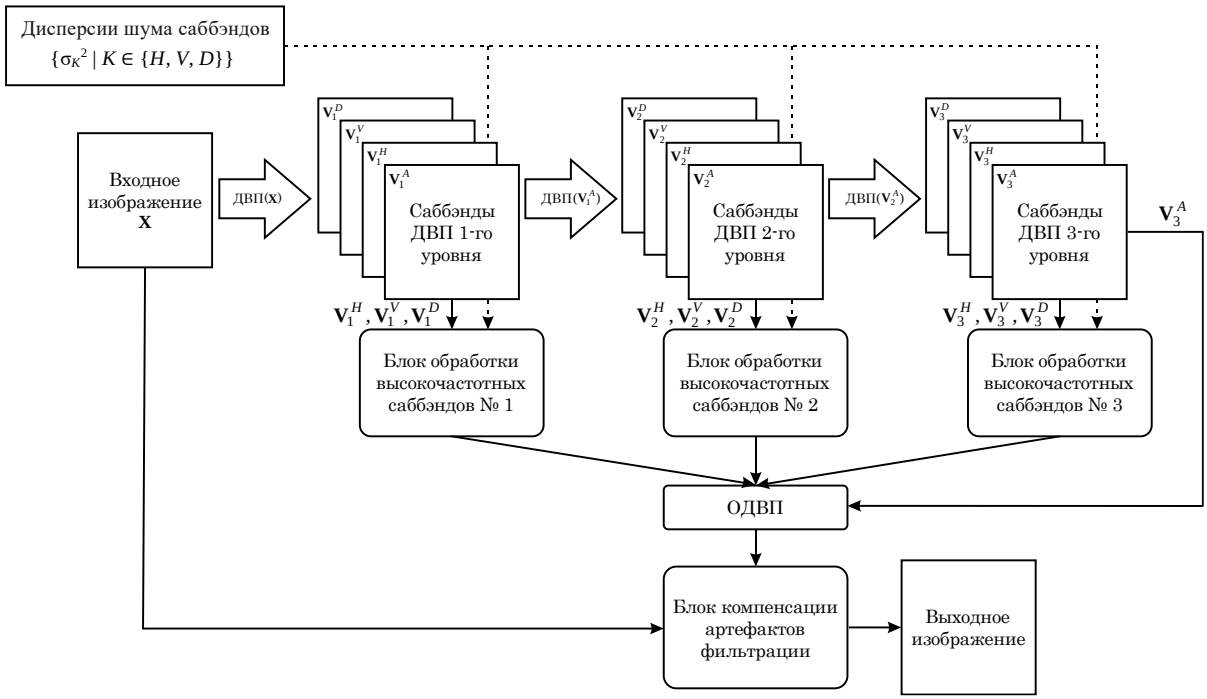


Рисунок 1 – Блок-схема архитектуры нейронной сети

Саббэнды $\mathbf{V}_i^H$, $\mathbf{V}_i^V$, $\mathbf{V}_i^D$ ($i \in \{1, 2, 3\}$) ДВП-спектра зашумлённого изображения вместе с СКО шума σ подаются на вход блокам обработки высокочастотных саббэндов, каждый из которых отвечает за фильтрацию вейвлет-коэффициентов одного из уровней ДВП. В процессе обработки высокочастотных саббэндов можно выделить 3 этапа:

I. Нечёткая классификация коэффициентов ДВП-спектра на «контурные» и «фоновые».
II. Получение оценок спектров мощности зашумлённого и полезного сигналов.
III. Вычисление коэффициентов фильтра Винера и подавление шума.

На первом этапе формируются маски $\mathbf{M}^K = \left(m_{k,l}\right)$, задающие для каждого вейвлет-коэффициента $v_{k,l} \in \mathbf{V}^K$ ($K \in \{H, V, D\}$) степень его принадлежности к контурной составляющей ДВП-спектра изображения как непрерывную величину $m_{k,l} \in [0; 1)$.

На втором этапе на основе найденных масок $\mathbf{M}^K$ выполняется расщепление ДВП-спектра на контурную и фоновую составляющие, для каждой из которых определяются оценки спектров мощности зашумлённого и полезного сигнала, необходимые для формирования коэффициентов фильтра Винера.

На третьем этапе вычисляются значения коэффициентов фильтра Винера для контурной и фоновой составляющих ДВП-спектра, после чего выполняется подавление шума путём покомпонентного перемножения матриц коэффициентов фильтра и вейвлет-коэффициентов саббэндов $\hat{\mathbf{U}}^K = \mathbf{W}^K \odot \mathbf{V}^K$. Так преобразованные саббэнды ДВП-спектра используются для получения восстановленного изображения с помощью ОДВП.

В процессе винеровской фильтрации вместе с шумовыми коэффициентами ДВП-спектра также могут быть подавлены вейвлет-коэффициенты, соответствующие полезной высокочастотной информации. Для того чтобы снизить влияние возникающих вследствие этого искажений и артефактов, в архитектуру модели дополнительно был добавлен блок (см. рисунок 1), разработанный на основе идей, изложенных авторами нейронной сети U-Net в работе [7]. Разработанный блок выполняет компенсацию артефактов, возникающих в процессе винеровской фильтрации; как показали проведённые нами эксперименты, его добавление даёт прирост пикового отношения сигнал/шум (PSNR) в среднем на 0,46 дБ по тестовым данным при среднем ($\sigma = 15$) и высоком ($\sigma = 30$) уровнях шума.

Обучение нейронной сети

Для обучения модели использовался набор данных ILSVRC [8], состоящий из 1 281 167 изображений. Исходные цветные изображения были преобразованы к полутоновому виду с разрядностью 8 бит/пиксель. В качестве эталонных меток в процессе обучения использовались фрагменты изображений размера 128×128 пикселей. Для формирования входных данных фрагмент-метка подвергался искажению белым аддитивным нормальным шумом с нулевым математическим ожиданием и СКО σ, значение которого генерировалось как реализация случайной величины, имеющей равномерное распределение $U(0, 30)$. Зашумлённые таким способом фрагменты изображений использовались затем в качестве образов обучающей выборки.

В качестве функции потерь использовался квадратичный штраф. Обучение модели выполнялось с помощью библиотеки PyTorch версии 2.4.1 [9]. При минимизации выборочного штрафа использовался алгоритм Adam [10] со значениями параметров: $\alpha = 10^{-2}$; $\beta_1 = 0{,}9$; $\beta_2 = 0{,}999$; $\varepsilon = 10^{-8}$, количество эпох было выбрано равным 600, размер батча выбран равным 8.

Тестирование алгоритмов

Тестирование сети выполнялось на наборе данных BSDS500 [11], состоящем из 500 полутоновых изображений. Оценка качества обработки выполнялась на основе сравнения значений мер PSNR и индекса структурного сходства (SSIM) [12] до и после обработки.

Проведено сравнение характеристик разработанной нейронной сети со следующими алгоритмами подавления шумов: винеровской фильтрацией в области ДВП [13] и винеровской фильтрацией с применением ансамбля нейросетевых моделей [5]. Тестовые данные формировались с трёмя фиксированными уровнями шума, определяемыми СКО: низкий уровень шума ($\sigma = 7$), средний ($\sigma = 15$) и высокий ($\sigma = 30$). Значения метрик качества, рассчитанные по результатам экспериментов и усреднённые по тестовому набору, представлены в таблице 1.

Таблица 1. Средние значения PSNR и SSIM до и после фильтрации при разных значениях СКО шума, вычисленные по набору BSDS500

Метрика	**PSNR (средн.) [дБ]**	**SSIM (средн.)**	**PSNR (средн.) [дБ]**	**SSIM (средн.)**	**PSNR (средн.) [дБ]**	**SSIM (средн.)**
	σ = 7		σ = 15		σ = 30	
Значение до обработки	31,26	0,80	24,75	0,56	18,94	0,32
Фильтр Винера [13]	34,59	*0,93*	29,84	0,83	25,93	0,67
Фильтр Винера + нейросетевой ансамбль [5]	*34,75*	*0,93*	*30,10*	*0,85*	*26,43*	*0,72*
Разработанный алгоритм	**35,20**	**0,94**	**31,17**	**0,87**	**27,49**	**0,76**

Примечание. Наибольшие значения в каждом столбце выделены жирным шрифтом, курсивом – следующие по убыванию после наибольших.

Полученные результаты позволяют сделать вывод, что предложенная нейронная сеть обеспечивает более качественное шумоподавление по сравнению с рассмотренными аналогами.

Применение фильтрации при распознавании зашумлённых изображений

Исследование влияния фильтрации на качество распознавания изображений, искажённых шумом, проводилось на основе идей, изложенных в работе [1]. Рассматривается задача детекции автомобильных номеров с применением нейросетевого детектора SqueezeDet [14]. Для сравнения влияния обработки различными алгоритмами фильтрации входных зашумлённых изображений на работу детектора проведены эксперименты с обучением SqueezeDet при использовании аугментаций 3-х типов:

– с добавлением шума в процессе обучения;
– с добавлением шума и последующей фильтрацией в процессе обучения;
– обучение без аугментации (с незашумлёнными дополнительно изображениями).

Для обучения детектора SqueezeDet также использовался алгоритм Adam. Наборы данных для обучения и тестирования, параметры обучения (величина шага градиентного спуска, размер батча, количество эпох и др.) выбраны такими же, как и в работе [1]. Сравнение качества работы детекторов, обученных с применением различных аугментаций, проводилось на основе анализа точности (доля верно распознанных номеров), доли номеров, распознанных с ошибками (хотя бы один символ номера распознан, при этом необязательно верно), и доли пропущенных номеров (ни один символ номера не обнаружен).

Результаты работы сетей, полученные на тестовом наборе изображений, зашумлённых при 3-х уровнях шума (σ ∈ {7; 15; 30}), представлены в таблице 2. Первый столбец таблицы содержит информацию об используемом типе аугментации и алгоритме подавления шумов на тестовом наборе данных: (I) – винеровская фильтрация в области ДВП [13], (II) – винеровская фильтрация с применением ансамбля нейросетвых моделей [5], (III) – предложенный в настоящей работе алгоритм. Прочерк "-" соответствует случаю, когда аугментация/фильтрация не использовались. Лучшие результаты в каждом столбце выделены жирным шрифтом.

Таблица 2. Результаты распознавания тестовых изображений при уровне шума σ ∈ {7; 15; 30}

Тип аугментации, алгоритм фильтрации	**Точность [%]**	**Доля номеров, распознанных с ошибками [%]**	**Доля пропущенных номеров [%]**
σ = 7			
-, -	80,74	12,1	7,16
Шум, -	91,72	5,54	**2,74**
Шум, фильтрация (I)	92,01	4,61	3,38
Шум, фильтрация (II)	92,01	4,61	3,38
Шум, фильтрация (III)	**92,66**	**4,25**	3,09

σ = 15			
-, -	52,63	21,63	25,74
Шум,-	91,83	5,4	**2,77**
Шум, фильтрация (I)	91,43	5,15	3,42
Шум, фильтрация (II)	91,14	5,4	3,46
Шум, фильтрация (III)	**92,30**	**4,57**	3,13
σ = 30			
-, -	12,17	19,73	68,1
Шум, -	88,44	8,82	**2,74**
Шум, фильтрация (I)	88,44	8,1	3,46
Шум, фильтрация (II)	89,27	**7,24**	3,49
Шум, фильтрация (III)	**89,31**	7,60	3,09

Результаты экспериментов демонстрируют, что обучение нейросетевого детектора с применением аугментации на основе разработанного алгоритма подавления шумов позволяет добиться наибольшей точности распознавания по сравнению с другими реализациями винеровской фильтрации. При среднем уровне шумов (σ = 15) предложенный в настоящем исследовании подход обеспечивает наибольший прирост доли верно распознанных номеров: на 1,16 % больше, чем при использовании нейросетевого ансамбля [5]. На рисунке 2 представлен пример зашумлённого изображения автомобильного номера, результат его фильтрации разработанным алгоритмом, а также результаты распознавания этих изображений нейросетевым детектором SqueezeDet.

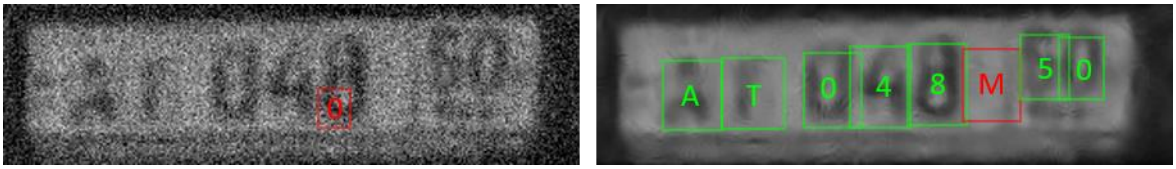


Рисунок 2 – Изображение номера “AT04850”, искажённого нормальным шумом с СКО σ = 30 (слева) и затем отфильтрованного (справа), и результаты распознавания сетью SqueezeDet (красным цветом выделены неверно детектированные символы, зелёным цветом – верно детектированные символы); значения метрик качества до/после обработки PSNR: 18,72/32,14 дБ, SSIM: 0,20/0,88

Также следует отметить, что архитектура предлагаемой нейронной сети содержит всего 1491 обучаемый параметр, что примерно в 1,8 раз меньше, чем у ансамбля из работы [5], который содержит 2772 обучаемых параметра, и на порядки меньше, чем у современных нейросетевых алгоритмов подавления шумов [2-4].

Заключение

Разработанный алгоритм нейросетевой фильтрации нормального белого шума обеспечивает наилучшие результаты шумоподавления в смысле роста метрик качества PSNR и SSIM, а также позволяет повысить точность распознавания (по методу [14]) изображений номерных знаков автомобилей по сравнению с рассмотренными альтернативами способов шумоподавления. Обладая при этом относительно малым количеством обучаемых параметров, предложенная нейронная сеть является более предпочтительным решением с точки зрения качества обработки и экономии вычислительных ресурсов.

Литература

1. Алимагадов К.А., Умняшкин С.В. Аугментация данных на основе вейвлет-фильтрации при обучении нейронных сетей // Труды Международной конференции по компьютерной графике и зрению "Графикон". - М.: Институт прикладной математики им. М.В. Келдыша РАН, 2023. - С. 437-442. - DOI 10.20948/graphicon-2023-437-442.

2. Huang J.-J., Dragotti P.L. WINNet: Wavelet-inspired invertible network for image denoising // IEEE Transactions on Image Processing. IEEE, 2022. - P. 4377-4392.

3. Herbreteau S., Kervrann C. DCT2net: An interpretable shallow CNN for image denoising // IEEE Trans. Image Process. IEEE, 2022. - P. 4292-4305.
4. Tu Z., Talebi H., Zhang H., Yang F., Milanfar P., Bovik A., Li Y. MAXIM: Multi-Axis MLP for image processing // IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2022. - P. 5759-5770.
5. Алимагадов К.А., Умняшкин С.В. Подавление белого шума на изображениях на основе винеровской фильтрации в области дискретного вейвлет-преобразования с применением нейросетевых технологий // Известия высших учебных заведений. Электроника. - 2022. - Т. 27, №6. - С. 807-818. - DOI 10.24151/1561-5405-2022-27-6-807-818.
6. Guo T. et al. A review of wavelet analysis and its applications: Challenges and opportunities // IEEE Access. - 2022. - №10. - P. 58869–58903. - DOI 10.1109/ACCESS.2022.3179517.
7. Ronneberger O., Fischer P., Brox T. U-Net: Convolutional networks for biomedical image segmentation // Medical Image Computing and Computer-Assisted Intervention. - Munich: Springer International Publishing, 2015. - P. 234-241.
8. Russakovsky O. et al. Imagenet large scale visual recognition challenge // International journal of computer vision. - 2015. - №115. - P. 211-252.
9. PyTorch // GitHub URL: https://github.com/pytorch/pytorch (дата обращения: 21.09.2024).
10. Kingma D.P., Ba J.L. Adam: a method for stochastic optimization // arXiv URL: https://arxiv.org/abs/1412.6980v1 (дата обращения: 21.09.2024).
11. Berkeley segmentation data set and benchmarks 500 (BSDS500) // Berkeley Computer Vision Group URL: https://www2.eecs.berkeley.edu/Research/Projects/CS/vision/grouping/resources.html (дата обращения: 15.09.2024).
12. Wang Z., Bovik A.C., Sheikh H.R., Simoncelli E.P. Image quality assessment: from error visibility to structural similarity // IEEE Transactions on Image Processing. - 2004. - №13. - P. 600–612.
13. Alimagadov K.A., Umnyashkin S.V. Application of Wiener filter to suppress white noise in images: wavelet vs Fourier basis // 2021 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (ElConRus). - St. Petersburg, Moscow: IEEE, 2021. - P. 2059-2063. – DOI 10.1109/ ElConRus51938.2021.9396470
14. Wu B. et al. Squeezedet: Unified, small, low power fully convolutional neural networks for real-time object detection for autonomous driving // Proceedings of the IEEE conference on computer vision and pattern recognition workshops, 2017. - P. 129–137.

NEURAL NETWORK WIENER FILTERING IN THE DOMAIN OF DISCRETE WAVELET TRANSFORM FOR IMPROVING QUALITY OF DIGITAL IMAGES RECOGNITION

PhD student Alimagadov K.A., Prof., Dr.Sci.Phys.-Math. Umnyashkin S.V.

National Research University of Electronic Technology MIET

The research is devoted to improvement of noise suppression in digital images distorted by additive white Gaussian noise. The neural network algorithm of Winer filtering in the domain of the discrete wavelet transform is developed. The estimation of noise suppression quality is based on comparison of peak signal to noise ratio (PSNR) and structural similarity index measure (SSIM) values. The proposed algorithm of noise suppression provides greater growth of quality metrics than the considered analogues. In addition, training neural network detector SqueezeDet with augmentations based on adding noise to input images with following filtering by the developed algorithm is considered for the task of license plate recognition. The experimental results show better accuracy for license plate recognition for the proposed algorithm of noise suppression applied for the preprocessing of input images before SqueezeDet detector.

СИНТЕЗ ЦВЕТОВЫХ ПРЕОБРАЗОВАНИЙ ДЛЯ ВИДЕОЭНДОСКОПИЧЕСКОЙ СИСТЕМЫ

доц. Мотыко А.А.[1,2], проф. Обухова Н.А.[1,2]

[1] Санкт-Петербургский государственный электротехнический университет «ЛЭТИ»
[2] Санкт-Петербургский государственный университет

Существующие подходы к синтезу цветовых преобразований, решая частные задачи и не учитывая специфику области применения, приводят к общей несогласованности обработки цветовой информации в эндоскопической видеосистеме. Предложен новый векторный критерий качества для синтеза цветовых преобразований, учитывающий как точность цветопередачи, так и передачу локального цветового контраста. Описаны соответствующие критерию методы синтеза цветовых преобразований на основе оптимизированных таблиц поиска с переменным разрешением. Рассмотрены примеры синтеза цветовых преобразований для целевых задач обработки эндоскопических изображений на основе предложенного критерия и методов.

Введение

Видеоэндоскопические системы сегодня играют ключевую роль в диагностике заболеваний внутренних органов. Независимо от области применения (бронхоскопия, гастроскопия, колоноскопия), телевизионная часть эндоскопа состоит из камеры и осветителя, расположенных на дистальном конце, а также видеопроцессора с устройством отображения. Видеоэндоскопия предполагает формирование данных в сложных специфических условиях, включающих использование широкоугольной оптики и точечного источника света, что приводит к геометрическим искажениям и неравномерному освещению кадра, наличию на изображении как очень ярких (из-за бликов), так и тёмных участков.

Цель обработки эндоскопических снимков – повышение их диагностической ценности для врача. На основе анализа снимка врач осуществляет диагностику, во многом ориентируясь на цветовую информацию. Поэтому важнейшими задачами обработки сигналов в видеоэндоскопах являются обеспечение точной цветопередачи и повышение локального цветового контраста, позволяющего выделить на изображениях целевые особенности. Последнее реализуется с помощью программных или аппаратных средств специальных режимов визуализации.

Несмотря на существование выраженной специфики в условиях получения изображений и в задачах их обработки, с точки зрения работы с цветовой информацией эндоскопы представляют собой типовые телевизионные системы, без учета особенностей области применения. Методы и алгоритмы синтеза цветовых преобразований не отличаются от применяемых в телевизионных системах общего назначения. Очевидно, что для максимизации пользы от обработки цветовой информации в эндоскопической системе необходимо, чтобы критерий и методы синтеза преобразований, осуществляющих обработку цвета, были бы ориентированы на целевой результат и учет специфики области применения. Разработке подобного критерия и методов, а также описанию результатов исследования их характеристик посвящена данная работа.

Векторный критерий качества цветовых преобразований для видеоэноскопии

Для обеспечения синтеза цветовых преобразований, учитывающего цель и задачи обработки эндоскопических снимков предложен векторный критерий качества, включающий как целевую функцию, отражающую точность цветопередачи, так и целевую функцию, отражающую уровень локального цветового контраста:

$$\mathbf{E}_{sum} = (\mathbf{E}_{color}(\mathbf{x}), E_{contrast}(\mathbf{x})),$$

где $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ – вектор параметров цветового преобразования; $E_{contrast}(\mathbf{x})$ – целевая функций, характеризующая передачу локального цветового контраста; $\mathbf{E}_{color}(\mathbf{x})$ – векторная целевая функция, характеризующая цветопередачу:

$$\mathbf{E}_{color}(\mathbf{x}) = \{f_{c1}(\mathbf{x}), f_{c2}(\mathbf{x}), ..., f_{cn}(\mathbf{x})\},$$

где $f_{c1}(\mathbf{x})$, $f_{c2}(\mathbf{x})$, …, $f_{cn}(\mathbf{x})$ – целевые функции, определяющие точность передачи отдельных цветов. Здесь терм, отвечающий за передачу локального цветового контраста, представляет собой скалярную функцию, соответствующую общей оценке локального контраста для изображения.

Предложенный векторный критерий дает возможность оценить параметры цветового преобразования путем решения оптимизационной задачи. Значение критерия в точке оптимума:

$$E^*_{sum} = \min_{\mathbf{x}}\{\mathbf{E}_{color}(\mathbf{x}), E_{contrast}(\mathbf{x})\}.$$

Соответствующее значение аргумента есть искомый вектор параметров преобразования:

$$\mathbf{x}_{opt} = \arg\min_{\mathbf{x}}\{\mathbf{E}_{sum}\}.$$

Важно отметить, что в существующих сегодня методах построения цветовых преобразований целевая функция, отвечающая за передачу контраста, отсутствует, а целевую функцию, отвечающую за цветопередачу, скаляризуют, устанавливая веса w_c для функций отдельных цветов равными. Фактически оценку параметров преобразования проводят по критерию минимизации средней или среднеквадратической ошибки цветопередачи. Провести скаляризацию для задачи оптимизации, согласно предложенному критерию, возможно лишь путем эмпирического подбора весов w_{color}, $w_{contrast}$, что не является простой задачей:

$$E^*_{opt} = \min_{\mathbf{x}}\{w_{color}E_{color}(\mathbf{x}) + w_{contrast}E_{contrast}(\mathbf{x})\}.$$

Для решения предлагается использовать известный метод, при котором одна из целевых функций сохраняется в качестве основной, а остальные переводятся в разряд ограничений. Таким образом, задачу можно переформулировать следующим образом:

$$E^*_{opt} = \min_{\mathbf{x}}\{E_{contrast}(\mathbf{x})\big|E_{color}(\mathbf{x}) \le \varepsilon\},$$

где ε – некоторый допустимый уровень для соответствующей целевой функции. Для оценки передачи контраста нужно использовать локальные меры, учитывающие цветовую информацию. Подходящими являются оценки на базе структурного тензора [3, 4].

В случае скаляризованной $E_{color}(\mathbf{x})$ такой уровень можно обоснованно определить. Предложено использовать известные пороги восприятия и приемлемости, выраженные в единицах CIEDE 2000 [2]. Считается, что если цветовое различие составляет менее 1.0-1.3 единицы, то человек не видит разницу между цветами, а при различии до 2 единиц разница ощущается, но воспринимается приемлемой. Тогда можно сформулировать оптимизационную задачу с ограничением, в качестве которого использовать среднюю или иную оценку ошибки цветопередачи (в совокупности с порогом приемлемости логично использовать оценку максимальной ошибки или 95%-й перцентиль). Варианты формулировки задачи:

$$E^*_{opt} = \min_{\mathbf{x}}\{E_{contrast}(\mathbf{x})\Big|\frac{1}{C}\sum_{c=1}^{C} w_c f_c(\mathbf{x}) \le 1.0\}.$$

$$E^*_{opt} = \min_{\mathbf{x}}\{E_{contrast}(\mathbf{x})\big|E_{color}(\mathbf{x})_{\max} \le \varepsilon \in [1..2]\},$$

$$E^*_{opt} = \min_{\mathbf{x}}\{E_{contrast}(\mathbf{x})\big|E_{color}(\mathbf{x})_{0.95} \le \varepsilon \in [1..2]\}.$$

Необходимость скаляризации целевых функций цветопередачи отдельных цветов в одну общую обусловлено применяемыми сегодня на практике методами оценки параметров цветовых преобразований и трудностями в определении значений весов для цветов или областей палитры. В результате при снижении значения какой-либо оценки средней ошибки, максимальные ошибки цветопередачи для отдельных цветов могут оставаться большими. Введение специальной регуляризации не устраняет в значительной мере проблему, поскольку оценка продолжает осуществляться по скаляризованному критерию, ориентированному на минимизацию средней ошибки. При этом именно максимальные ошибки наиболее заметны наблюдателю и снижают как субъективно воспринимаемое качество изображения, так и точность диагностики. Следовательно, в рамках синтезированного векторного критерия необходимо разработать метод оценки параметров цветовых преобразований, ориентированный не только на минимизацию средней ошибки цветопередачи, но и, учитывая контекст или важность отдельных участков цветовой палитры для диагностики, на минимизацию максимальной ошибки для этих цветов.

Метод синтеза цветовых преобразований для видеоэндоскопической системы

Существует множество моделей для цветовых преобразований: регрессионные линейные и полиномиальные модели (в том числе т.н. рут-полиномиальные [5]), модели на основе сплайнов [6], нейронных сетей [7] и трехмерных таблиц поиска [8]. Известны аналитические способы оценки параметров цветовых преобразований [5]. В ряде случаев для решения используют

численные методы минимизации функционалов. При этом, если задействуются локальные методы, то остро встает вопрос с выбором точки начального приближения [9]. В публикации [10] авторами настоящей работы был предложен метод численной оценки параметров цветовых преобразований, учитывающий свойства зрения.

В задачах цветовых преобразований широко используются трехмерные таблицы поиска [11]. Таблица характеризуется своей размерностью – числом индексов (входных переменных), которое задает нужное значение функции и числом узлов решетки, то есть точек, в которых предварительно рассчитано значение аппроксимируемой функции. Например, N=4 обозначает трехмерную таблицу размером 4x4x4 узлов. Таблицу формализуют как набор **V** узлов **a** в пространстве прообраза $\mathbb{R}^n$ и соответствующих им значений **b** в пространстве образа $\mathbb{R}^m$:

$$\mathbf{V} = [\mathbf{a}_i \in \mathbb{R}^n, \mathbf{b}_i \in \mathbb{R}^m].$$

Расположение узлов $\mathbf{a}_i$ разделяет n-мерную область g на небольшие подобласти, а соответствующие значения $\mathbf{b}_i$ являются приблизительными оценками отображения g в точках $\mathbf{a}_i$. С учетом интерполяции, значения вычисляют следующим образом. Определяют веса w_{ij} для отображаемой точки $\mathbf{c}_i$ в пространстве прообраза:

$$\sum_{j=1}^{N_v} w_{ij}\mathbf{a}_j = \mathbf{c}_i, \sum_{j=1}^{N_v} w_{ij} = 1,$$

где N_v – число узлов решетки таблицы. Эти веса используют для получения оценки значения преобразования (то есть координаты в пространстве образа) в точке $\mathbf{c}_i$:

$$\widehat{\mathbf{d}}_i = \sum_{j=1}^{N_v} w_{ij}\mathbf{b}_j.$$

Таблицы поиска не лишены недостатков [10]. Поэтому развивают оптимизированные таблицы [12], в которых положения узлов в пространстве прообраза и их значений в пространстве образа рассчитаны во взаимосвязи посредством некоторой процедуры, обеспечивающей лучшую точность отображения. Обычно построение оптимизированных таблиц опирается на «обучающий» набор данных N_T точек $\mathbf{c}_i$ в пространстве прообраза, для которых значения в образе $\mathbf{d}_i$ известны. Важным является использование при построении таблиц регуляризаторов [13], направленных на придание преобразованию различных свойств – гладкости, невыхода за пределы и пр. Существующие сегодня методы построения оптимизированных таблиц для цветовых преобразований не соответствуют требованиям, поставленным в рамках синтезированного векторного критерия – они, как и другие методы, ориентированы на минимизацию средней ошибки цветопередачи. При этом присущее таблицам поиска разбиение отображаемых пространств на отдельные участки соответствует идее векторного представления целевых функций цветопередачи. В данной работе авторами предлагается метод оценки параметров цветовых преобразований, соответствующий введенному критерию и позволяющий учитывать не только среднюю ошибку, но и максимальные ошибки для важных цветов и областей палитры.

В работе [14] авторами был предложен аналитический метод построения оптимизированных таблиц поиска. Полный функционал, который нужно минимизировать в рамках построения оптимизированной таблицы поиска, осуществляющей цветовое преобразование, имеет вид:

$$E(\mathbf{b}) = \mathbf{b}^T\left(\mathbf{F} + \lambda_S\mathbf{S} + \lambda_K\mathbf{K} + \lambda_\rho\mathbf{H}\right)\mathbf{b} - 2\left(\mathbf{g} - \lambda_\rho\mathbf{s}\right)^T\mathbf{b} + h + \lambda_\rho t,$$

где **F** (симметричная квадратная матрица), **g** (вектор столбец), h (свободный член) – составляющие части квадратичной формы, состоящей из координат точек N_T и весов интерполяции, отражающей ошибку цветопередачи; **S**, **s**, **K**, **H**, t – составляющие квадратичных форм, отвечающих за регуляризацию различного назначения [17]; λ_S, λ_K, λ_ρ – веса слагаемых регуляризации; **b** – искомые координаты узлов оптимизированной таблицы поиска в пространстве образа. Решение задачи минимизации находится в явном виде:

$$\mathbf{b}_{opt} = \left(\mathbf{F} + \lambda_S\mathbf{S} + \lambda_K\mathbf{K} + \lambda_\rho\mathbf{H}\right)^{-1}\left(\mathbf{g} - \lambda_\rho\mathbf{s}\right).$$

При построении оптимизированных таблиц поиска с помощью данного метода [14], существует возможность локализации ошибок выполняемого отображения в соответствующих ячейках таблицы. На основе этого наблюдения предложен метод построения оптимизированных таблиц с переменным разрешением. Таблица строится в рамках двухпроходного алгоритма. На

первом этапе строится базовая оптимизированная таблица поиска. После проводится анализ, в каких ячейках базовой таблицы ошибка преобразования (рассчитанная по обучающему набору) превышает заданное (допустимое) значение. В этих ячейках осуществляется увеличение разрешения таблицы. Заданное значение может иметь смысл допустимой максимальной ошибки преобразования, при превышении которой необходимо увеличить разрешение таблицы в этой области для повышения точности.

С точки зрения предложенного векторного критерия задание областей таблицы с повышенным разрешением имеет смысл автоматического определения отдельных участков палитры, в которых фиксируется максимальная ошибка преобразования, и придания им большего веса при расчете параметров преобразования. Процесс локального повышения разрешения можно представить как вставку патчей в базовую таблицу (рисунок 1).

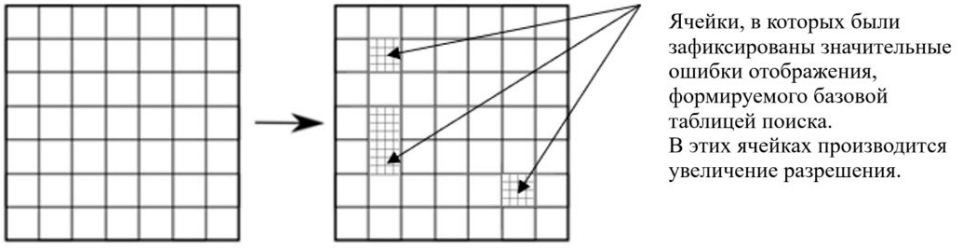


Рисунок 1 – Принцип оптимизированной таблицы поиска с переменным разрешением

Преимуществом является то, что базовая таблица, и патчи представляют собой оптимизированные таблицы, построенные по одному методу. Для каждого патча параметры разрешения и регуляризации выбираются независимо от базовой таблицы, а построение происходит на подвыборке обучающего набора, входящего в соответствующую ячейку базовой таблицы. Результирующая таблица поиска обладает неоднородными свойствами (по своему объему), задаваемыми локальными регуляризаторами. Это позволяет с высокой точностью локально адаптировать таблицу с переменным разрешением в соответствии с рельефом аппроксимируемого преобразования, чтобы адекватно описать имеющиеся нелинейности.

Применение предложенного векторного критерия качества и методов синтеза цветовых преобразований в задачах обработки эндоскопических изображений

Предложенный векторный критерий и метод синтеза цветовых преобразований были успешно применены в ряде задач видеоэндоскопии. Получены решения, для режимов специальной визуализации, обеспечивающих повышение локального цветового контраста [15, 16], для задачи синтеза единого изображения в бимодальной эндоскопии [17]. В данной работе приведены результаты сравнительного исследования точности цветопередачи в решении задачи характеризации сенсора при разработке видеоэндоскопа, обеспечиваемой цветовыми преобразованиями, полученными различными методами.

Задачи характеризации сенсора и баланса белого являются базовыми при создании любой телевизионной системы, в том числе видеоэндоскопа. При разработке устройства могут быть использованы различные сенсоры, для которых необходимо выполнять отображение цветовых пространств из зависимого от устройства (как правило RGB) в независимое (обычно XYZ). Особенностью является то, что формирование изображения видеоэндоскопом происходит с единственным источником освещения, что делает процедуру баланса белого простой. В рамках описываемой разработки использовался сенсор IMX715. При создании видеоэндоскопических систем используют фильтр ИК отсечки. Он позволяет подавить чувствительность синего и зеленого каналов в области «красных» длин волн, приводящую к падению локального цветового контраста при визуализации сосудистых структур. Характеристики используемого фильтра (BG-40) и сенсора приведены на рисунке 2. В работе использовался источник освещения, характеристики которого практически соответствовали типу D65. В качестве исходных данных для построения преобразований (набора пар цветов в двух пространствах) методом спектральной реконструкции [18] была сформирована сетка размером 40x40x40 точек с учетом имеющихся известных характеристик сенсора, фильтра и источника освещения. В качестве данных для тестирования использовались колорчекеры: Munsell extended, DC и набор цветовых градиентов.

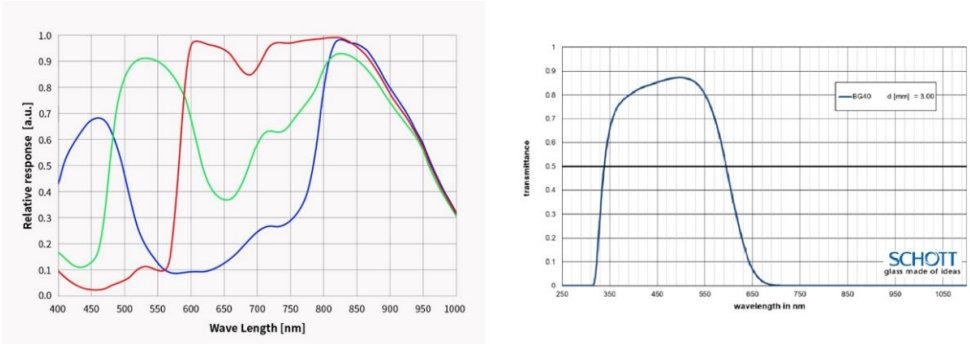


Рисунок 2 – Спектральные характеристики сенсора IMX715 (слева), фильтра BG-40 (справа)

Было проведено сравнительное исследование методов и моделей синтеза цветовых преобразований в рамках решаемой задачи: линейное матричное 3x3 и рут-полиномиальное 3x6 преобразования [5], коэффициенты получены по исходным данным методом наименьших квадратов; неоптимизированная таблица поиска N=17, построенная на базе спектральной сетки; оптимизированная таблица поиска размером N=17 и оптимизированная таблица с переменным разрешением, построенные с помощью предложенного метода по исходным данным.

Линейное матричное преобразование 3x3 является на сегодня наиболее используемым в реальных устройствах аналитическим цветовым преобразованием. Рут-полиномиальное преобразование 3x6 представляет собой перспективную замену матричному цветокорректору и на данный момент внедряется в передовые устройства. Неоптимизированная таблица размером N=17 – базовый и часто встречающийся вариант, применяемый в современной технике для задач цветовых преобразований. Два остальных метода предложены в рамках данной работы. При построении оптимизированной таблицы был выбран размер N=17, исходя из объема обучающего набора. Для версии с переменным разрешением повышение разрешения происходило в ячейках, где максимальная ошибка по обучающему набору превышала значение ΔE2000 = 3. Таких ячеек в рамках используемой выборки оказалось две. В них были встроены патчи размером N=4. В таблицу 1 сведены усредненные результаты для методов, полученные по всем тестам.

Таблица 1. Усредненные результаты для методов, полученные по всем тестам

Метод	средняя ошибка, ΔE2000	90%-й перцентиль, ΔE2000	95%-й перцентиль, ΔE2000	максимальная ошибка, ΔE2000
линейное 3x3	1,32	2,59	4,34	6,81
рут-полин. 3x6	1,14	2,34	3,23	5,57
неопт. LUT N=17	0,93	2,59	3,43	13,59
опт. LUT N=17	0,54	1,94	2,38	4,94
опт. LUT N=17 с перем. разреш.	0,47	1,80	2,19	3,77

Исходя из анализа таблицы, можно сделать следующие выводы. Реализация оптимизированной таблицы N=17 с переменным разрешением (встроено 2 патча N=4) позволила обеспечить наилучшие точностные характеристики. Средняя ошибка меньше в 1.98 – 2.81 раза, 90%-й перцентиль в 1.3 – 1.44 раза, 95%-й перцентиль в 1.47-1.98 раза, максимальная ошибка в 1.48 – 3.6 раз. Важным является значительное снижение предложенным методом максимальных ошибок цветового преобразования (включая 95%-й перцентиль), наиболее важных с практической точки зрения. Полученные результаты подтверждают корректность и практическую значимость предложенного критерия и соответствующего ему метода синтеза цветовых преобразований.

Источник финансирования

Исследование (раздел “Применение предложенного векторного критерия качества и методов синтеза цветовых преобразований в задачах обработки эндоскопических изображений”) выполнено при финансовой поддержке Министерства науки и высшего образования Российской Федерации в рамках реализации научного проекта по соглашению № 075-15-2025-013

Литература

1. Брахман Т.Р. Многокритериальность и выбор альтернативы в технике. М.: Радио и связь, 1984, 287 с.
2. Kim A-Ri, Kim Hong-Suk, Park Seung-ok, Measuring of the Perceptibility and Acceptability in Various Color Quality Measures, Journal of the Optical Society of Korea, 2011, v.15.
3. Chatoux H. и др., Gradient in spectral and color images: from the Di Zenzo initial construction to a generic proposition, J. Opt. Soc. Am. A, 2019, v. 36, № 11, p. 154
4. Di Zenzo S. A note on the gradient of a multi-image, Comput. Vision, Graph. Image Process, 1986. v. 33, № 1. pp. 116–125.
5. Zhao Y., Ferguson S., Zhou H., Elliott C., Rafferty K. Color alignment for relative color constancy via non-standard references. IEEE Transactions on image processing, 2022, v31, p.6591-6604
6. Suominen J. Camera color correction using splines. Electronic Imaging, 2024, pp. 1651 – 1656
7. Wu P.-Ch., Lin C. A green and practical color photograph printing technology based on color difference model and human perception. IEEE Access, 2022, no. 10, pp. 649 – 666.
8. Zeng H., Cai J., Li L., Cao Z., Zhang L. Learning image-adaptive 3D lookup tables for high performance photo enhancement in real-time. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022, no. 4, pp. 2058-2073, doi: 10.1109/TPAMI.2020.3026740
9. Mishra, S.K., Panda, G., et al. On q-BFGS algorithm for unconstrained optimization problems. Advances in Difference Equations, vol. 1, p.1-24, 2020. https://doi.org/10.1186/s13662-020-03100-2
10. Мотыко А.А., Обухова Н.А. Метод перцептуальной цветокоррекции // Информационно-управляющие системы. 2023. № 3. С. 26-38.
11. Mezhenin A.V., Pauk M.A. Reduction of visual artifacts when using three-dimensional color conversion tables taking into account the extended dynamic range. Culture and Technologies, 2023. v. 8, no. 4, pp. 204-211 (In Russian). doi:10.17586/2587-800X-2023-8-4-204-211
12. Srivastava S., Delp E., Ha T., Allebach J. Color Management Using Optimal Three-Dimensional Look-Up Tables. Journal of Imaging Science and Technology, 2010, №54, p. 030402-1 – 030402-14.
13. Garcia E., Gupta M. Optimized construction of icc profiles by lattice regression. Final Program and Proceedings - IS and T/SID Color Imaging Conference. 2010, pp. 353-358.
14. Мотыко А. А., Обухова Н. А., Якубович Ю. В. Метод синтеза оптимизированных таблиц поиска для цветовых преобразований. Информационно-управляющие системы, 2025, № 1, с. xx
15. N. A. Obukhova, A. A. Motyko, A. A. Pozdeev. The Endoscopic Images Visualization in Clinical Decision Support Systems, 22th International Conference on Digital Signal Processing and its Applications (DSPA), Moscow, Russia, 2020, pp. 1-4, doi: 10.1109/DSPA48919.2020.9213285.
16. Обухова Н.А., Мотыко А.А., Поздеев А.А. Исследование и разработка методов улучшения эндоскопических (медицинских) изображений. Известия высших учебных заведений России. Радиоэлектроника. 2019; 22(2):22-30. https://doi.org/10.32603/1993-8985-2018-22-2-22-30
17. Обухова Н.А., Мотыко А.А., Поздеев А.А., Савельева О.Р., Ян С. Цветовые преобразования видеоэндоскопических изображений. Материалы 34-й Международной конференции по компьютерной графике и машинному зрению GraphiCon. Омск, 2024, с. 641-655.
18. Yi-Tun Lin, Graham D. Finlayson. A Rehabilitation of Pixel-Based Spectral Reconstruction from RGB Images. Sensors, 2023, no. 8, pp. 1-20, doi: 10.3390/s23084155

SYNTHESIS OF COLOR TRANSFORMATIONS FOR VIDEOENDOSCOPIC SYSTEM

Motyko A.A.[1,2], Obukhova N.A.[1,2]

[1] Saint Petersburg Electrotechnical University "LETI"
[2] Saint Petersburg State University

Existing approaches to the synthesis of color transformations, solving individual problems and not considering the specifics of the application area, lead to a general inconsistency of color information processing in the endoscopic video system. A new quality criterion for the synthesis of color transformations is proposed, considering both the accuracy of color reproduction and the transmission of local color contrast. Methods of color transformation synthesis based on optimized lookup tables with variable resolution corresponding to the criterion are described. Examples of synthesis of color transformations for target tasks of endoscopic image processing based on the proposed criterion and methods are described.

СИСТЕМА СЛИЯНИЯ МРТ И КТ-СНИМКОВ БРЮШНОЙ ПОЛОСТИ НА ОСНОВЕ АНАТОМИЧЕСКОГО ВНИМАНИЯ

асп. Нишчхал Н.[1]

[1] Сибирский государственный университет науки и технологий имени академика М. Ф. Решетнева, г. Красноярск, Россия

В последние годы визуализация медицинских снимков стала одним из важнейших инструментов в диагностике широкого спектра заболеваний. Слияние изображений магнитно-резонансной томографии (МРТ) и компьютерной томографии (КТ) открывает новые возможности для повышения точности диагностики. Однако процесс слияния изображений требует тщательного учета анатомических особенностей исследуемой области, что является сложной задачей. В работе предлагается новый подход к слиянию изображений МРТ и КТ на основе механизма анатомического внимания. Разработка фокусируется на выделении ключевых анатомических структур, что обеспечивает сохранение целостности визуализации органов брюшной полости и повышение качества итоговых изображений. Эксперименты на базе данных изображений брюшной полости подтверждают эффективность предлагаемого подхода, демонстрируя значительные улучшения по сравнению с традиционными методами.

Введение

Современные методы медицинской визуализации важны для диагностики заболеваний брюшной полости. МРТ чувствительна к мягким тканям, подходит для оценки печени, поджелудочной железы и почек. КТ обладает высоким разрешением и точностью определения плотности тканей, полезна для анализа костей и локализации органов. Объединение данных МРТ и КТ повышает точность диагностики, но сталкивается с трудностями, связанными с различиями в получении данных, утратой анатомической точности и ограниченной интеграцией характеристик.

СЛияние изображений МРТ и КТ приобретает особую значимость в клинических сценариях, требующих высокой степени точности, таких как обнаружение малых новообразований, дифференциальная диагностика тканей и предоперационная подготовка. Для успешной реализации этих задач необходимо разработать методику, которая способна поддерживать анатомическую целостность, подчеркивать ключевые зоны каждой модальности и минимизировать возникновение артефактов. Настоящее исследование посвящено разработке системы слияния МРТ и КТ изображений, основанной на механизме анатомического внимания. Целью данного подхода является сохранение структурной целостности и максимизация информативной ценности синтезированных изображений. Для этого в процессе слияния применяются анатомические приоритеты, извлечённые из предварительно обученных моделей сегментации, что позволяет акцентировать внимание на важных структурах, таких как печень, почки и поджелудочная железа. Многослойная нейронная сеть используется для интеграции информации на различных уровнях детализации, тогда как функции потерь, учитывающие точность реконструкции, структурную целостность и визуальную качество, способствуют эффективному обучению модели.

Предлагаемый метод направлен на преодоление ограничений существующих подходов к слиянию изображений, открывая новые возможности для повышения точности диагностики и эффективности терапевтических вмешательств при заболеваниях брюшной полости.

Обзор методов слияния медицинских изображений

На начальных этапах использовались вейвлет-преобразования, которые позволяли анализировать изображения на уровне частотных компонент. Эти методы продемонстрировали хорошие результаты при слиянии МРТ и КТ изображений, но часто приводили к потере анатомических деталей [1, 2]. Другие подходы, основанные на плотности градиента, улучшали сохранение краевых областей, однако имели ограничения в интеграции высокоуровневых признаков [3].

С развитием глубокого обучения произошел значительный прорыв. Свёрточные нейронные сети (СНС) стали основой современных решений, улучшая сохранение структурных деталей, но требовали больших объемов аннотированных данных для обучения [4]. Более сложные

архитектуры использовали плотные соединения для эффективного объединения признаков, хотя не всегда учитывали анатомическую значимость различных частей изображения [5].

Перспективным направлением является использование механизмов внимания, таких как модель Attention U-Net, которая позволяет системе фокусироваться на ключевых частях изображения [6]. Исследования показали, что применение масок органов для управления процессом объединения может значительно повысить клиническую ценность результирующих изображений [7].

Современные методы объединения медицинских изображений сталкиваются с потерей мелких деталей и требуют больших объемов размеченных данных для глубоких нейросетей, что затрудняет их применение в клинической практике. Недостаток внимания к анатомическим структурам снижает эффективность этих методов.

Предлагаемая методика анатомического внимания устраняет существующие недостатки, сочетая преимущества механизмов внимания и сегментации для точного контроля процесса объединения. Применение многослойной СНС с учётом анатомического внимания и значимости структурных компонентов не только повышает качество изображения, но и усиливает его клиническое значение.

Методология

В основе предложенного метода лежит модель AbdomenNet [8], специально разработанная глубокая нейросетевая архитектура, предварительно обученная на обширной базе медицинских изображений брюшной полости. Данная модель оптимизирована под задачи мультиклассовой сегментации и демонстрирует высокие показатели точности, достигая коэффициента сходства Dice, превышающего 0,85, для следующих анатомических структур: печень, почки, поджелудочная железа, желчевыводящая система (включая желчный пузырь и основные желчные протоки) и селезенка (с выделением полного объема органа). Результатом работы модели AbdomenNet являются высокоточные анатомические маски с разрешением 512×512 пикселов, которые служат пространственными ориентирами для дальнейших этапов обработки. Эти маски обеспечивают точную локализацию ключевых анатомических компонентов, акцентируя внимание на значимых областях изображения и обеспечивая основу для последующей интеграции данных. На рисунке 1 показаны маски ключевых анатомических областей, где красный, белый, зелёный, синий и жёлтый цвета обозначают, соответственно, печень, позвоночный отдел, почку, селезёнку и поджелудочную железу.

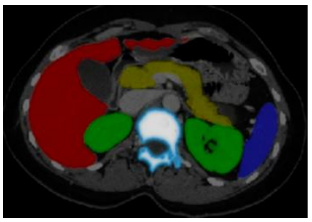


Рисунок 1 – Маски ключевых анатомических областей

Модуль внимания представляет собой ключевой компонент, способствующий улучшению процесса интеграции мультимодальной информации посредством выделения наиболее значимых регионов изображения на основании их анатомического значения. Данная система реализована в виде многослойной архитектуры, включающей глобальную сеть внимания и набор локальных подсистем, которые обеспечивают детализированное обнаружение целевых областей.

Глобальная сеть внимания функционирует на уровне полного разрешения входных изображений (512×512), выполняя задачу обнаружения взаимосвязей между различными модальностями (например, МРТ и КТ). Она генерирует карты внимания, основываясь на пересечениях и корреляционных связях между характеристиками двух изображений, при этом каждому пикселю присваивается весовой коэффициент $\alpha(x, y) \in [0,1]$, который отражает степень его значимости в общем контексте. Этот метод позволяет идентифицировать крупные анатомические структуры, обеспечивая базис для последующего глубокого анализа на локальном уровне.

Локальные подсистемы внимания направлены на обработку отдельных зон интереса, таких как конкретные органы и их границы. Карты внимания создаются отдельно для каждой целевой области, что способствует более точному определению ключевых анатомических компонентов. Высшие весовые коэффициенты присваиваются следующим регионам:

– границы органов, где градиент интенсивности превышает установленный порог (> 0.6);

– области интересов, такие как зоны с потенциальными патологическими изменениями;

– важные анатомические ориентиры, играющие существенную роль в клинической интерпретации.

С целью повышения надежности и точности карты внимания проходят процесс итеративной оптимизации, состоящий из трех этапов. На каждом этапе весовые коэффициенты корректируются с учетом вновь выявленных характеристик, что позволяет улучшить точность карт внимания и минимизировать влияние шумов или несущественных деталей. Таким образом, механизм внимания помогает модели фокусироваться на значимых частях данных, улучшая качество изображений и обеспечивая анатомическую точность. информации.

Сеть объединения реализует иерархический подход к интеграции признаков из МРТ и КТ изображений, обеспечивая сохранение анатомической точности и высокого качества итогового изображения. Архитектура сети состоит из трёх основных этапов: извлечение признаков, интеграция признаков и генерация результирующего изображения.

На этапе извлечения признаков входные изображения из обеих модальностей (МРТ и КТ) обрабатываются через параллельные потоки. Для каждого потока используются многоуровневые свёрточные слои с ядрами различного размера (3×3 и 5×5), что позволяет выделять как мелкомасштабные, так и крупномасштабные особенности. Каждому свёрточному слою сопутствует нормализация пакета (Batch Normalization) для стабилизации обучения и функция активации ReLU для добавления нелинейности. Это обеспечивает эффективное извлечение пространственных и текстурных признаков из каждого изображения.

На этапе интеграции признаков используется стратегия взвешенного объединения признаков с учетом карт внимания, рассчитанных на предыдущем этапе. Для каждого пикселя результирующей карты признаков интеграция выполняется по формуле:

$$F_{\text{fused}}(x,y) = \alpha(x,y) \cdot F_{\text{MRI}}(x,y) + (1 - \alpha(x,y)) \cdot F_{\text{CT}}(x,y), \quad (1)$$

где F_{MRI} и F_{CT} – карты признаков из МРТ и КТ соответственно, а α(x, y) – веса, полученные из карт внимания. Этот подход позволяет сохранить уникальные характеристики обоих модальностей, акцентируя внимание на ключевых областях, таких как границы органов и патологически измененные участки. Кроме того, архитектура включает пропускные соединения (skip connections), которые передают исходные признаки из начальных слоев к более глубоким слоям сети. Это предотвращает потерю пространственной информации и улучшает согласованность результатов.

На стадии генерации итогового изображения осуществляется многослойная реконструкция данных на различных масштабных уровнях (1/4, ½ и полном разрешении). Данная стратегия гарантирует точное восстановление деталей на каждом уровне. Для улучшения качества изображения применяется остаточное обучение (residual learning), которое позволяет нейросети учитывать различия между входными и целевыми значениями, что делает изображение более детальным. Конечное изображение генерируется с разрешением 512×512 пикселов, сохраняющее как общую анатомическую структуру, так и мелкие текстурные детали.

Таким образом, сеть объединения успешно интегрирует информацию из двух модальностей, обеспечивая высокую информативность и визуальную достоверность синтезированных изображений, что делает ее важным инструментом для диагностики.

Для достижения точного и качественного объединения изображений используются три ключевые функции потерь: потери восстановления, потери анатомической точности и потери визуального качества

Потери восстановления (Reconstruction Loss) минимизирует разницу между результирующим изображением (I_{fused}) исходными изображениями (I_{MR} и I_{CT}):

$$L_{\text{recon}} = \| I_{\text{fused}} - I_{\text{MR}} \|^2 + \| I_{\text{fused}} - I_{\text{CT}} \|^2, \quad (2)$$

Потери анатомической точности (Anatomical Accuracy Loss) оценивает перекрытие сегментированных масок результирующего изображения и эталона с помощью Dice Loss:

$$L_{\text{anat}} = 1 - \frac{2|M_{\text{fused}} \cap M_{\text{truth}}|}{|M_{\text{fused}}| + |M_{\text{truth}}|}, \quad (3)$$

где M_{fused} – сегментированная маска результирующего изображения, M_{truth} – эталонная сегментированная маска, $|M|$ – площадь (количество пикселов) маски.

Потери визуального качества (Perceptual Loss) сравнивает высокоуровневые признаки результирующего изображения с эталонными, извлеченные из предварительно обученной сети.

Итоговая функция потерь – это объединение всех компонент с весами λ_{recon}, λ_{anat}, λ_{perc}.

$$L_{\text{total}} = \lambda_{\text{recon}} L_{\text{recon}} + \lambda_{\text{anat}} L_{\text{anat}} + \lambda_{\text{perc}} L_{\text{perc}} \quad (4)$$

Экспериментальные результаты

В экспериментальной части работы была проведена проверка эффективности предложенного метода слияния МРТ и КТ изображений посредством экспериментов с реальными клиническими данными. В этом разделе приводятся подробности эксперимента, используемые наборы данных и результаты тестирования. Для целей обучения и оценки использовался набор данных TCIA. Он содержит МРТ и КТ-изображения брюшной полости, включающая данные о 100 пациентах с разнообразными патологиями органов брюшной полости, а также аннотированные маски ключевых анатомических структур (таких как печень, почки, поджелудочная железа и др.), а также сегментированные маски, полученные посредством применения предварительно обученной модели сегментации AbdomenNet, демонстрирующей высокий уровень точности сегментации (*Dice Similarity Coefficient* > 0.85). В процессе обучения модель использовала 80% данных, тогда как оставшаяся часть (20%) была зарезервирована для этапа тестирования. Предварительная обработка данных включала регистрацию изображений, нормализацию интенсивности и выделение областей интереса. Обучение модели осуществлялось с применением функционалов потерь, учитывающих восстановление, анатомическую точность и перцепционную составляющую. Обучающий процесс проводился на графическом процессоре NVIDIA A100 при размере пакета 8 и количестве эпох 100.

В процессе решения задачи объединения МРТ и КТ изображений брюшной полости акцент делается на выбор методик, обеспечивающих высокую степень точности при сохранении анатомической структуры и улучшении визуального качества. Для количественной оценки эффективности различных подходов применяются следующие объективные показатели:

– среднеквадратическое отклонение (*SD*) отражает общую дисперсию контрастности изображения, что является критическим параметром для чёткой визуализации границ органов;

– энтропия (*EN*) оценивает информационное содержание объединённого изображения, играя важную роль в выявлении мелких анатомических особенностей;

– нормированная взаимная информация (Q_{MI}) измеряет степень сохранения информации из исходных изображений после их объединения;

– фазовая когерентность (Q_{PC}) демонстрирует уровень сохранности текстур и границ органов в результирующем изображении.

Результаты экспериментальных исследований подтвердили преимущество разработанного подхода с модулем анатомической внимания по сравнению с классическими методами по нескольким ключевым параметрам. В частности, объединенные изображения, полученные с использованием предложенного алгоритма, демонстрируют повышенную контрастность, лучшее сохранение анатомических контуров и снижение потерь информации. В таблице 1 приведен сравнительный анализ результатов.

Таблица 1. Сравнительный анализ

Модель	***SD***	***EN***	Q_{MI}	Q_{PC}
DCT [9]	47,6423	5,5367	0,7513	0,7074
PCNN [10]	48,5357	5,5842	0,7647	0,7382
DCNN [11]	50,7705	5,8216	0,8208	0,7523
Предложенный метод	55,3834	6,1733	0,8424	0,7861

Полученные результаты подтверждают эффективность предлагаемого метода, который демонстрирует лучшее сохранение структуры органов и текстуры по сравнению с существующими подходами. Высокие значения *SD*, *EN*, Q_{MI} и Q_{PC} подчеркивают его пригодность для клинического использования, где критически важны точность и качество визуализации.

Для примера визуального анализа была выбрана поджелудочная железа, поражённая опухолью, с соответствующими КТ и МРТ изображениями. Изображения были объединены четырьмя различными методами: предложенного метода с анатомическим вниманием, а также моделей PCNN, DCT и DCNN. После объединения поджелудочная железа была сегментирована на всех четырёх объединённых изображениях с помощью модели AbdomenNet. Этот процесс обеспечивал более чёткую идентификацию поражённых областей, что позволило лучше оценить степень распространения опухоли. Результаты показали, что предложенный метод превзошёл другие подходы, формируя изображения с повышенной структурной чёткостью и высокой точностью сегментации. Эти наблюдения подчеркивают эффективность предложенной методики для повышения диагностической точности в сложных клинических ситуациях.

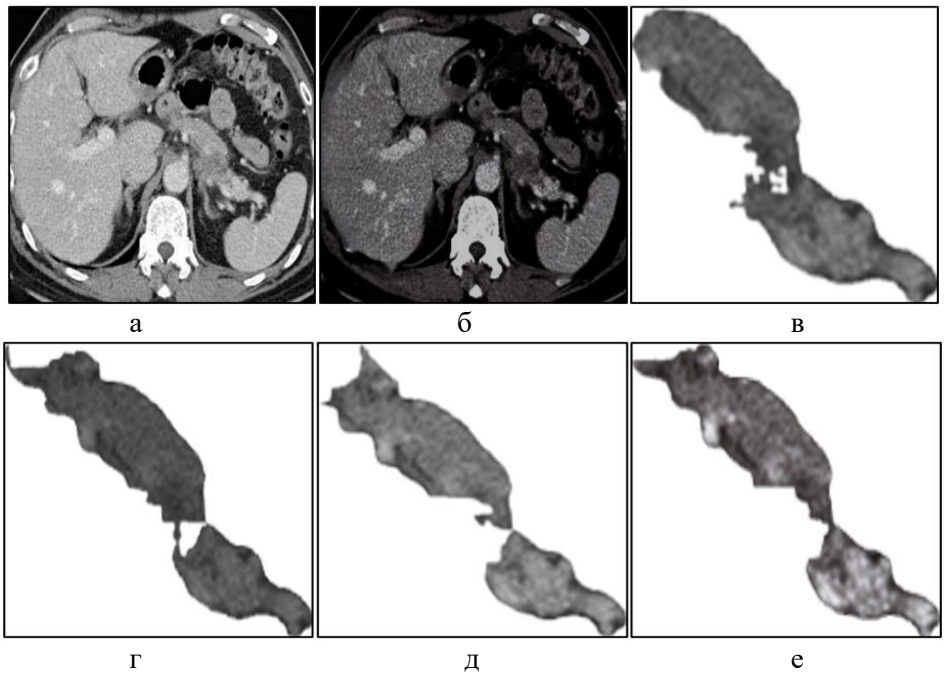


Рисунок 2 – Сравнение визуальных результатов: **а** исходное КТ-изображение, **б** исходное МРТ-изображение, **в** сегментированная поджелудочная железа из изображения, объединенного методом DCT, **г** сегментированная поджелудочная железа из изображения, объединенного методом PCNN, **д** сегментированная поджелудочная железа из изображения, объединенного методом DCNN, **е** сегментированная поджелудочная железа из изображения, объединенного предложенным методом.

Заключение

В данной работе предложен анатомически обоснованный метод слияния МРТ и КТ изображений органов брюшной полости. Благодаря интеграции анатомических знаний в процедуру объединения изображений, метод демонстрирует значительное улучшение в сохранении анатомической целостности и повышении диагностической значимости. Данный подход представляет собой перспективное решение для мультидисциплинарной медицинской визуализации с высоким потенциалом клинического применения.

Литература

1. Tawfik, N., Elnemr, H.A., Fakhr, M. et al. Survey study of multimodality medical image fusion methods. Multimed Tools Appl 80, 6369–6396 (2021).
2. Azam, M.A., Khan, K.B., Ahmad, M., Mazzara, M. Multi modal medical image registration and fusion for quality enhancement. Comput Mater Continua 68, 821–840 (2021).
3. Thakur, N.V., Yenurkar, G.K., Aherrao, A., Landge, S., Katre, S. Medical image fusion using discrete wavelet transform: In view of deep learning. 2023 1st DMIHER International Conference on Artificial Intelligence in Education and Industry 4.0 (IDICAIEI), Wardha, India, 2023, pp. 1-6.
4. Liang, N. Medical image fusion with deep neural networks. Sci Rep 14, 7972 (2024).
5. Liu, Y., Chen, X., Ward, R.K., Wang, Z.J. Medical image fusion via convolutional sparsity based morphological component analysis. IEEE Signal Process. Lett. 26, 485–489 (2019).
6. Oktay, O., Schlemper, J., Le Folgoc, L. et al. Attention U-Net: Learning where to look for the pancreas. Medical Image Computing and Computer-Assisted Intervention (MICCAI), 2018, 3-12.
7. Liu, J., Mao, S., Pan, L. Attention-based two-branch hybrid fusion network for medical image segmentation. Applied Sciences 14(10), 4073 (2024).
8. Rickmann, A.M., Senapati, J., Kovalenko, O., Peters, A., Bamberg, F., Wachinger, C. AbdomenNet: deep neural network for abdominal organ segmentation in epidemiologic imaging studies. BMC Med Imaging 22(1), 168 (2022).
9. Chen, J. Digital image progressive fusion method based on discrete cosine transform. Journal of Mathematics, vol. 2023 (2023).
10. Liu, L., Huo, J. PCNN Model Guided by Saliency Mechanism for Image Fusion in Transform Domain. Sensors (Basel) 23(5), 2488 (2023).
11. Yousif, A.S., Omar, Z., Sheikh, U.U., Khalid, S.A. A new scheme of medical image fusion using deep convolutional neural network and local energy pixel domain. 2020 IEEE-EMBS Conference on Biomedical Engineering and Sciences (IECBES), Langkawi Island, Malaysia, 2021, pp. 384-388.

ANATOMICAL ATTENTION SUPERVISED MR-CT IMAGE FUSION FOR ABDOMINAL IMAGING

PhD stud. Nishchhal N.[1]

[1] 1Reshetnev Siberian State University of Science and Technology, Institute of Informatics and Telecommunications, 31, Krasnoyarsky Rabochy ave., Krasnoyarsk, 660037 Russian Federation

Abdominal imaging benefits from the complementary strengths of magnetic resonance (MR) and computed tomography (CT) modalities. While MR offers superior soft-tissue contrast, CT provides high-resolution bone and organ structure visualization. Image fusion combines these modalities to improve diagnostic accuracy, particularly in abdominal conditions. This paper introduces an anatomical attention-based supervised framework for MR-CT image fusion. The proposed method embeds anatomical priors into an attention mechanism, selectively emphasizing relevant features from each modality. Results demonstrate significant improvements in preserving anatomical details and enhancing clinical utility, validated by quantitative metrics and qualitative analysis.

ОБНАРУЖЕНИЕ КУПОЛА СЛЕПОЙ КИШКИ НА ВИДЕОЭНДОСКОПИЧЕСКИХ ИЗОБРАЖЕНИЯХ

доц. Хрящев В.В., асп. Котов Н.В., Тихомиров А.А.

Ярославский государственный университет им. П.Г. Демидова

Разработан и исследован алгоритм на основе методов глубокого обучения для автоматического обнаружения купола слепой кишки на видеоизображениях колоноскопического исследования кишечника. Для тестирования использована база видеоданных 43 пациентов с колоноскопических исследований. Разработанный комбинированный алгоритм АДКТ–ОП демонстрирует высокую точность обнаружения купола слепой кишки, что подтверждается метрикой F1 на уровне 0,89. Внедрение предложенного алгоритма позволит усилить контроль качества колоноскопических исследований.

Введение

Современные технологии анализа видеоизображений находят широкое применение в медицинской диагностике, особенно в эндоскопической диагностике желудочно-кишечного тракта. За последние 10 лет благодаря развитию нейросетевых технологий и глубокого обучения достигнуты значительные успехи в разработке алгоритмов детектирования, сегментации и классификации медицинских изображений. Эти методы позволяют улучшить качество диагностики, повысить производительность врачей-эндоскопистов и контролировать качество процедур [1-3].

Подавляющее большинство исследований в области анализа колоноскопических видеоизображений посвящено задаче обнаружения полипов кишечника. Такое исследование является актуальным по следующим причинам [4-7]:

- Полипы в толстой кишке могут со временем перерождаться в злокачественные опухоли. Своевременное обнаружение и удаление полипов значительно снижает риск развития колоректального рака, который является одним из самых распространенных онкологических заболеваний.
- Человеческий фактор может привести к пропуску полипов, особенно мелких или плоских. Автоматизированный анализ изображений с использованием алгоритмов глубокого машинного обучения помогает повысить точность диагностики, уменьшая вероятность ошибок.
- Использование искусственного интеллекта для анализа изображений способствует стандартизации диагностики, что особенно важно в условиях разного уровня квалификации специалистов.
- Внедрение технологий анализа изображений позволяет повысить качество медицинской помощи, особенно в регионах с ограниченным доступом к высококвалифицированным специалистам.
- Автоматизированный анализ позволяет собирать и анализировать большие объемы видеоданных, что может быть полезно для дальнейших исследований в области гастроэнтерологии и онкологии.

Целью данной работы является разработка алгоритма для автоматического обнаружения купола слепой кишки на видеоданных колоноскопического исследования. Достижение купола слепой кишки считается одним из критериев полноценной колоноскопии. Слепая кишка — это начальный отдел толстой кишки, и достижение её купола подтверждает, что врач осмотрел всю толстую кишку. Это важно для исключения патологий, которые могут локализоваться в этом отделе, включая полипы, опухоли или воспалительные изменения. Если этот этап не выполнен, обследование считается неполным, что может привести к пропуску патологий. Во многих странах достижение купола слепой кишки является обязательным требованием для проведения качественной колоноскопии. Это подтверждается визуализацией анатомических ориентиров, таких как илеоцекальный клапан и аппендикулярное отверстие. Однако в некоторых случаях достижение купола слепой кишки может быть затруднено из-за анатомических особенностей пациента. В таких ситуациях врач должен документально зафиксировать причины неполного обследования и рассмотреть возможность альтернативных методов диагностики, таких как

виртуальная колоноскопия (с использованием компьютерной томографии) или капсульная эндоскопия.

Эта работа является продолжением исследований, проводимых коллективом авторов из Ярославского государственного университета и Ярославской областной клинической онкологической больницы [8-9].

Описание набора данных

Для проведения компьютерных экспериментов была создана база из колоноскопических исследований 43 пациентов с видеозаписями области купола слепой кишки. Видеоматериалы были обработаны и размечены врачами-эндоскопистами. Собранная база имеет следующие основные характеристики:

- Средняя длительность видео: 17 минут 28 секунд.
- Частота кадров в секунду: 25.
- Разрешение: 720×576 пикселей.
- Среднее время нахождения области купола на видеозаписи: 1 минута 19 секунд.

Примеры изображений купола слепой кишки представлены на рисунке 1.



а) б) в)

Рисунок 1 – Примеры изображений купола слепой кишки

Алгоритмы обнаружения области купола

Из-за анатомических особенностей кишечника купол слепой кишки может быть трудно различим от складок или бугров на стенке кишечника. Предложенный подход включает нейросетевой детектор, алгоритм трекинга объектов и дополнительный модуль на основе анализа оптического потока. Для обнаружения купола слепой кишки в видеопотоке использовался нейросетевой детектор YOLOv8. Однако точность детектирования страдала от большого числа ложных срабатываний из-за анатомической схожести пространства кишечника [8-9]. Для повышения точности в процесс был внедрен алгоритм трекинга (АДКТ) на базе алгоритма ByteTrack [10]. АДКТ использует информацию о соседних кадрах для фильтрации ложных срабатываний.

Для улучшения работы детектора дополнительно был добавлен модуль для обработки оптического потока между видеокадрами, основанный на алгоритме Хорна-Шанка [11]. Этот модуль позволяет определить направление движения эндоскопа и контролировать качество колоноскопических исследований. Точка изменения направления вычисляется как среднее арифметическое предсказанных направлений за 20 секунд: движение вперед — (+1), остановка — (0), движение назад — (-1). Для обучения классификатора использовались размеченные видеоданные, связывающие карты смещений с направлением движения эндоскопа.

Методы вычисления оптического потока рассчитывают движение между кадрами видеопоследовательности, взятыми в моменты t и $t+\Delta t$ в позиции каждого пикселя. Эти методы базируются на разложении функции, описывающей сигнал изображений $I(x, y, t)$, в трехмерный ряд Тейлора с сохранением только слагаемых первого порядка:

$$I\,(x+\delta x, y+\delta y, t+\delta t) = I\,(x,y,t) + \delta x \frac{\partial I}{\partial x}(x,y,t) + +\delta y \frac{\partial I}{\partial y}(x,y,t)\ \delta t \frac{\partial I}{\partial t}(x,y,t). \quad (1)$$

В качестве δt берется разница между моментами съемки двух последовательных кадров: $I\,(x,y,t)$ и $I\,(x,y,t+1)$. С помощью δx и δy оценивается перемещение пикселя в интервале времени от t до $t+1$. Предполагаем, что сигнал изображения остается постоянным при бесконечно малых сдвигах по трехмерным координатам, т.е.

$I\,(x+\delta x, y+\delta y, t+\delta t) = I\,(x,y,t)$, тогда выражение (1) принимает следующий вид:

$$0 = \delta x \frac{\partial I}{\partial x}(x,y,t) + \delta y \frac{\partial I}{\partial y}(x,y,t) + \delta t \frac{\partial I}{\partial t}(x,y,t). \quad (2)$$

Разделив на δt получим:

$$0 = \frac{\delta x}{\delta t}\frac{\partial I}{\partial x}(x,y,t) + \frac{\delta y}{\delta t}\frac{\partial I}{\partial y}(x,y,t) + \frac{\partial I}{\partial t}(x,y,t). \quad (3)$$

Его можно записать в более компактном виде:

$$I_x V_x + I_y V_y + I_t = 0, \quad (4)$$

где I_x, I_y, I_t – производные функции сигнала изображения при (x,y,t) в соответствующих направлениях, V_x и V_y – компоненты оптического потока для $I\,(x,y,t)$. Для решения уравнения оптического потока (4) часто используется итерационная схема. При этом в оригинальном алгоритме Хорна-Шанка [11] используются следующие аппроксимации для частных производных I_x, I_y, I_t:

$$I_x(x,y,t) = \frac{1}{4}[I(x+1,y,t) + I(x+1,y,t+1) + I(x+1,y+1,t) + I(x+1,y+1,t+1)] - \frac{1}{4}[I(x,y,t) + I(x,y,t+1) + I(x,y+1,t) + I(x,y+1,t+1)], \quad (5)$$

$$I_y(x,y,t) = \frac{1}{4}[I(x,y+1,t) + I(x\,y+1,t+1) + I(x+1,y+1,t) + I(x+1,y+1,t+1)] - \frac{1}{4}[I(x,y,t) + I(x,y,t+1) + I(x+1,y,t) + I(x+1,y,t+1)], \quad (6)$$

$$I_t(x,y,t) = \frac{1}{4}[I(x,y,t+1) + I(x,y+1,t+1) + I(x+1,y,t+1) + I(x+1,y+1,t+1)] - \frac{1}{4}[I(x,y,t) + I(x,y+1,t) + I(x+1,y,t) + I(x+1,y+1,t)]. \quad (7)$$

Итоговый алгоритм (АДКТ-ОП) объединяет алгоритмы оптического потока и детектирования с трекингом, результаты которых классифицируются дополнительным модулем. Он определяет временной интервал присутствия купола или выдает сообщение об его отсутствии, сигнализируя о возможном нарушении регламента проведения колоноскопии.

Основой метрикой оценки точности работы алгоритмов на основе детектора является F1-мера или среднее гармоническое точности (P) и полноты (R):

$$\mathrm{F1} = 2\frac{P * R}{P + R}. \quad (8)$$

Комбинированный алгоритм АДКТ-ОП демонстрирует высокую точность – метрика F1 достигает значения 0,89.

Заключение

Для обучения и тестирования нейронной сети в Ярославской областной онкологической

больнице врачи-эндоскописты провели разметку видеоматериала. На основе этой разметки была создана база видеоизображений из 43 колоноскопических обследований.

Основной целью исследования была модернизация алгоритма детектирования купола слепой кишки. Внедрение алгоритмов трекинга и оценки оптического потока улучшило точность по метрике F1 после добавления нейросетевого классификатора.

Разработанный алгоритм подходит для постобработки архивных данных и использования в реальном времени во время колоноскопии. Это позволит объективно контролировать качество обследования в соответствии с медицинскими стандартами.

Литература

1. Гудфеллоу Я., Бенджио И., Курвилль А. Глубокое обучение // ДМК-Пресс. 2017. 652 с.
2. Zhou S.K., Greenspan H., and Shen D. Deep Learning for Medical Image Analysis. Elsevier Science (2017).
3. Munzer B., Schoeffmann K., Boszormenyi L. Content-based processing and analysis of endoscopic images and videos: A survey. Multimed. Tools Appl. (2018) 77. pp. 1323–1362.
4. Кашин С.В., Никонов Е.Л., Нехайкова Н.В., Лилеев Д.В. Стандарты качественной колоноскопии (пособие для врачей). Доказательная гастроэнтерология. 2019; 8(1-2):3–32.
5. Akbar M., Shima R., Soroushmehr S.M.R., Karimi N., Samavi S., and Najarian K. Classification of informative frames in colonoscopy videos using convolutional neural networks with binarized weights // International conference of the IEEE Engineering in Medicine and Biology Society. 2018. pp. 65–68.
6. Ballesteros C., Trujillo M., Mazo C., Chaves D., Hoyos J. 2016. Automatic classification of non-informative frames in colonoscopy videos using texture analysis. In: Iberoamerican congress on pattern recognition. Cham: Springer, 401–408.
7. Bernstein C., Thorn M., Monsees K., Spell R., O'Connor J.B. 2005. A prospective study of factors that determine cecal intubation time at colonoscopy. Gastrointestinal Endoscopy 61(1):72–75.
8. Хрящев В.В., Завьялов Д.В., Андержанова А.С. Классификация эндоскопических изображений устья червеобразного отростка на основе методов глубокого машинного обучения // Цифровая обработка сигналов. 2023. № 1. С. 35–38.
9. Котов Н.В., Хрящев В.В., Тихомиров А.А. Исследование нейросетевой архитектуры YOLO для обнаружения изображения купола слепой кишки в колоноскопии // Цифровая обработка сигналов и ее применение (DSPA-2024): докл. 26-й междунар. конф. Москва, 2024. С. 151–155.
10. Zhang Y., Sun P., Jiang Y., Yu D., Weng F., Yuan Z., Luo P., Liu W., Wang X. Bytetrack: Multi-object tracking by associating every detection box // European Conference on Computer Vision. Springer 2022. pp. 1-21.
11. Horn B., Schunck B. Determining optic flow. Artificial intelligence, vol. 17, 1981, pp. 185–203.

DETECTION OF THE CECULE DOME IN VIDEO ENDOSCOPIC IMAGES

Khryashchev V.V., Kotov N.V., Tikhomirov A.A.

P.G. Demidov Yaroslavl State University

An algorithm based on deep learning methods for automatic detection of the cecum dome in video images of colonoscopic examination of the intestine was developed and studied. A video database of 43 patients with colonoscopic examinations was used for testing. The developed combined ADCT-OP algorithm demonstrates high accuracy of detection of the cecum dome, which is confirmed by the F1 metric at the level of 0.89. The implementation of the proposed algorithm will improve the quality control of colonoscopic examinations.

ПРИМЕНЕНИЕ МЕТОДОВ КОМПЬЮТЕРНОГО ЗРЕНИЯ И МАШИННОГО ОБУЧЕНИЯ ДЛЯ СКРИНИНГА ЗАБОЛЕВАНИЙ МОЛОЧНОЙ ЖЕЛЕЗЫ ПО ТЕРМОМАММОГРАФИЧЕСКИМ СНИМКАМ

д.т.н. Хамухин А.В.[1], д.ф.-м.н., проф. Умняшкин С.В.[1], д.м.н., проф. Долгов И.М.[2],
студ. Комольцева Д.В.[1], студ. Ларев А.А.[1]

[1] Национальный исследовательский университет «МИЭТ»
[2] ООО «Дигносис»

Предложен метод автоматизированного анализа термографических снимков молочных желез на предмет наличия злокачественных, доброкачественных образований или их отсутствия. Метод основан на сегментации изображений с помощью нейросетевой модели и обработке температурных признаков на полученных сегментах. В ходе исследования проанализированы возможности использования различных признаков описания термоизображения, проведена оценка их влияния на качество распознавания категорий пациентов. В качестве базового способа построения классификатора использовался метод опорных векторов.

Результаты экспериментов показали, что учет локальных температурных аномалий и анализ свойств симметрии распределения температуры в молочных железах без учета сосудистых областей позволяет достичь доли правильной классификации в 86%. Разработанный метод подтверждает эффективность использования машинного обучения в сочетании с термографией как одного из инструментов скрининга рака молочной железы.

Введение

Рак молочной железы (МЖ) – одно из наиболее часто диагностируемых онкологических заболеваний у женщин. По данным статьи Всемирной организации здравоохранения [1], в 2022 году рак молочной железы был диагностирован у 2,3 миллиона женщин и вызвал 670 000 случаев смерти во всем мире.

Раннее выявление рака молочной железы повышает выживаемость, снижает агрессивность терапии и улучшает качество жизни пациента [2, 3]. Однако скрининг на ранних стадиях заболевания остается сложным и дорогостоящим мероприятием, в силу чего разработка дополнительного неинвазивного метода диагностики является актуальной задачей.

Обзор источников показывает растущий интерес к неинвазивным методам диагностики, основанным на анализе термографических снимков с применением компьютерного зрения и машинного обучения. Например, в статье [4] представлена система компьютерной диагностики для обнаружения рака молочной железы с использованием термографических изображений и алгоритмов машинного обучения. Авторы применили тепловизионную камеру FLIR One Gen 3, подключаемую к смартфонам, что делает метод более доступным. Система автоматизирует процесс классификации термограмм, достигая высокой доли правильных ответов алгоритма (98,83%), полноты (100%) и точности (98,44%). В работах [5] и [6] рассматриваются методы классификации термографических снимков с использованием метода опорных векторов на основе признаков, характеризующих температурные аномалии молочных желез, симметрии температур между правой и левой молочными железами, текстурных особенностей термографических изображений. Приведенные работы показывают эффективность, преимущества по сравнению с традиционными способами идентификации рака и потенциал применения компьютерных систем для диагностирования заболевания по термомаммографическим снимкам.

Главные проблемы, с которыми сталкиваются исследователи в рассматриваемой области – это выбор способов получения снимков, предобработки и сегментации изображений, а также выбор признаков, на основе которых строится решающее правило, определяющее наличие или отсутствие заболеваний. Для проведения эффективной сегментации снимков мы обучили и применили нейросеть, которая является развитием подхода YOLO [7, 8]. В отличие от других работ в данной области, это позволило проводить локальный анализ термоизображений именно в области МЖ. Помимо часто используемых признаков, таких как характеристики тепловых аномалий, текстурные описания и характеристики температурной симметрии при сравнении левой и правой МЖ, в нашей работе рассматриваются новые подходы к выявлению особенностей термоснимков: средняя разность и корреляция гистограмм температур в областях квадрантов, выделенных с помощью нейросети сегментов; моменты первого и второго порядков для наиболее горячих и

наиболее холодных областей, их ориентация и геометрия; различие температур между квадрантами МЖ без учетов наиболее горячих сосудов и другие.

Разработка алгоритма классификации термоснимков

Разработанный нами метод термомаммографии включает семь этапов: получение изображений, предобработка изображений, выделение областей МЖ, разделение областей МЖ на квадранты, выделение признаков, отбор признаков, классификация снимков по 3 классам: снимки без отклонений (класс «норма»), снимки пациентов с доброкачественными новообразованиями (класс «ДН»), снимки с злокачественными новообразованиям (класс «ЗН»).

1. Получение изображений

Исходные тепловизионные изображения с разрешением 640×480 пикселей предоставлены ООО «Дигносис» (Россия). Термоснимки были получены с использованием тепловизора «ТВС300-мед» [9] производства ООО «СИЛАР» (Россия) и комплекса медицинский программной обработки и анализа термограмм «TVision», созданного в компании «Дигносис».
Исходный набор данных, предоставленный для анализа, составляет 298 серий снимков для каждого пациента. Каждая серия содержит три термограммы, в пикселях которых хранится информация о температуре, зафиксированной тепловизором с точностью до 0,03 °С (см. рисунок 1). Такая серия снимков в трёх проекциях (передней, правой косой, левой косой) создаётся в помещении с необходимым температурным режимом.

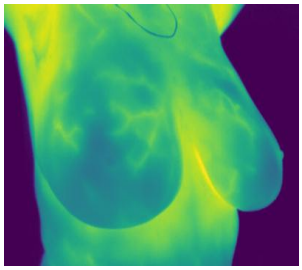
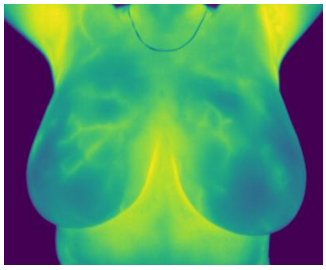
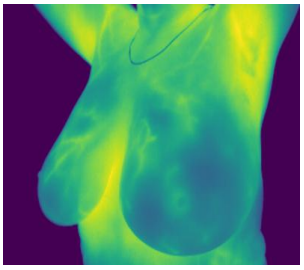
а  б  в 

Рисунок 1 – пример термограмм: а – правая косая, б – прямая, в – левая косая

2. Предобработка изображений

Предварительная обработка изображений включает в себя изменение размера, нормализацию, подавление шума, повышение контрастности и т.д. В нашем исследовании предварительная обработка проводилась в 2 этапа: приведение к единому размеру изображения 320×240 точек и приведение диапазона температур изображения к единому для всех экземпляров значению [25, 40] °С, чтобы исключить влияние температуры окружающей среды.

3. Выделение областей МЖ

Чтобы эффективно работать с областью МЖ и не учитывать другие контрастные области, такие как зоны подмышек, в рамках данной работы обучена нейросеть, которая выполняет сегментацию термоснимков, проставляя по 10 точек контуров на каждую МЖ слева и справа. Для поиска данных точек подготовлен обучающий набор данных из нескольких сотен серий снимков с разметкой в виде точек с индексами от 1 до 10 слева и справа, см. рисунок 2.

Подготовка нейронной сети, осуществляющей сегментацию, проведена с помощью методов, описанных в работах [8, 10-12] и реализованных в библиотеке Ultralytics [7]. В результате получен нейросетевой алгоритм, с достаточной высокой точностью предсказывающий положение точек контура изображений МЖ для переданных на вход термоснимков, в том числе не входящих в выборку обучения.

4. Разделение областей МЖ на квадранты

Разделение на квадранты является важным этапом в разработанном методе. Описания снимков УЗИ врачей-маммологов для МЖ часто ссылаются на аномалии в том или ином «квадранте» области МЖ (см., например, [13]), которые принято обозначать «верхне-наружный», «верхне-внутренний», «нижне-наружный» и «нижне-внутренний». Приняв такое поквадрантное представление МЖ, нами был разработан новый алгоритм автоматического предсказания положения данных областей на снимке.

Первый этап алгоритма заключается в запуске нейросетевой процедуры поиска точек контуров МЖ с индексами от 1 до 10, см. рисунок 2. Далее используются середины некоторых хорд области МЖ, а именно середины отрезков точек с индексами i, $11 - i$, где индекс i является натуральным числом от 1 до 5 (на рисунке 2 точки отмечены жёлтым цветом). Через полученные средние точки по методу наименьших квадратов проводится условно «вертикальная» прямая, а через центр масс точек контура перпендикулярно «вертикальной» прямой проводится условно «горизонтальная» прямая (на рисунке 2 прямые отмечены синим цветом).

В конечном итоге результатом работы алгоритма являются маски каждого квадранта МЖ, ограниченного найденными прямыми и контуром МЖ, который можно приблизить многоугольником с вершинами из найденной нейросетевым алгоритмом 10-точечной последовательности, см. рисунок 2.

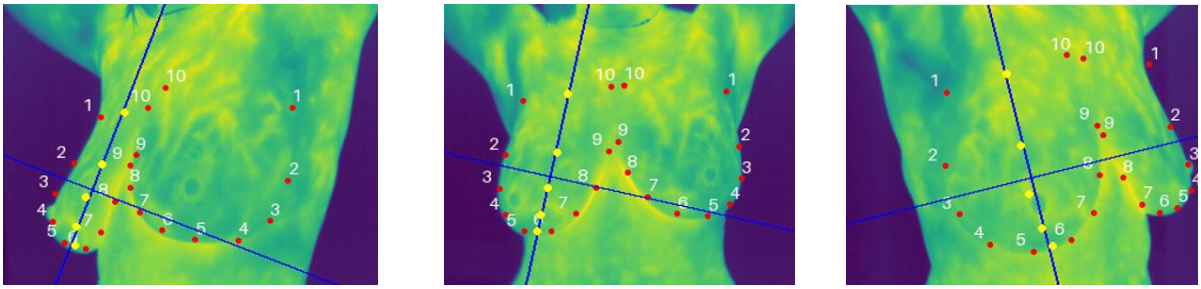


Рисунок 2 – пример контурных точек, расставленных нейросетью (красные) и прямых, делящих область МЖ на квадранты

5. Выделение признаков

Ключевым этапом данного исследования является вычисление признаков, часто используемых для классификации новообразований в МЖ в исследованиях со смежной тематикой, и разработка оригинальных характеристик, описывающих различные тепловые особенности тканей МЖ. К последним, ранее не встречаемым нами в литературе, новым признакам относятся, например, вейвлет-подобные температурные разности в автоматически выделяемых квадрантах, текстурные признаки, признаки характеристик форм экстремальных температурных областей. Далее приводится описание использованных и разработанных характеристик и способов их расчёта.

Средняя разность между квадрантами и корреляция гистограмм температур по квадрантам. Проведенное на предыдущем этапе разделение на квадранты позволяет изучить распределение температур в разных областях как для каждой МЖ, так и провести сравнение между левой и правой МЖ. Мы использовали для этого следующие признаки: средняя разность и корреляция гистограмм распределения температур между квадрантами.

Статистические признаки. Коэффициенты эксцесса и асимметрии распределений температур, рассчитанные для каждой МЖ.

Текстурные признаки на основе анализа матрицы GLCM и признаков Харалика. Из матрицы смежности уровней серого (Gray Level Co-Occurrence Matrix, GLCM) извлекаются признаки Харалика для различных направлений (0°, 45°, 90° и 135°). Использование этих характеристик позволяет учитывать ориентированность текстур в тепловой структуре тканей МЖ.

Линейное преобразование Хафа для обнаружения сосудистых линий. Преобразование Хафа позволяет выделять прямые линии на термографических изображениях МЖ, которые описывают сосудистую структуру. Среднее значение и энтропия аккумуляторной матрицы Хафа помогают оценить однородность сосудистой сети и выявить аномалии, такие как утолщение или увеличение количества сосудов.

Признаки симметричности. Разделение МЖ на квадранты позволяет сравнивать температурные распределения в симметрично расположенных областях левой и правой МЖ, различия в их температурных характеристиках в численном выражении позволяет выявить асимметрии, которые могут сигнализировать о наличии патологического состояния.

Температурные признаки для проекций молочных желез. В исследовании температурных характеристик МЖ, полученных в разных проекциях (прямой и боковых), проводилось сопоставление распределения температуры для одной и той же железы, отображаемой в различных ракурсах. Таким образом, левая МЖ, изображённая на прямой проекции, сравнивалась с левой МЖ,

представленной на правой боковой проекции, а правая МЖ – с изображением на левой боковой проекции.

Гистограммные признаки. Для извлечения гистограммных признаков предварительно были рассчитаны усреднённые несмещённые гистограммы температурного распределения для каждого класса. Затем определялось расстояние между гистограммами температурного распределения анализируемого изображения в прямых проекциях и заранее вычисленными эталонными гистограммами, принадлежащих к серии изображений с известными категориями («норма», «ДН», «ЗН»). Чем ближе полученное распределение к одной из эталонных гистограмм, тем выше вероятность принадлежности пациента к соответствующему классу, а метрика близости гистограмм является одним из числовых признаков для решающего правила.

Характеристики тензора инерции температурных областей в окрестности локальных экстремумов. Для вычисления данных признаков разработан алгоритм со следующим псевдокодом.

Шаг 1. Определить множество точек D как множество точек МЖ, выделенных с помощью вершин полигона нейросетевым алгоритмом.

Шаг 2. Положить номер области $r = 1$.

Шаг 3. Выполнить поиск точки (m_x^r, m_y^r) с максимальным значением температуры T_m^r на множестве D.

Шаг 4. С помощью алгоритма выделения связных компонент («flood fill») выделить область связанных точек R^r, начиная с точки (m_x^r, m_y^r), такую, что в любой точке области $(i,j) \in R^r$ её температура $I_{ij} \in [(1-\Delta)T_m^r, (1+\Delta)T_m^r]$, где $\Delta = 0{,}05$ – небольшое значение.

Шаг 5. Для области R^r вычислить моменты первого и второго порядка
$M_x^r = \sum_{(i,j)\in R^r} i \cdot I_{ij}$, $M_y^r = \sum_{(i,j)\in R^r} j \cdot I_{ij}$, $M_{xx}^r = \sum_{(i,j)\in R^r}(i-x_c)^2 I_{ij}$,
$M_{yy}^r = \sum_{(i,j)\in R^r}(j-y_c)^2 I_{ij}$, $M_{xy}^r = \sum_{(i,j)\in R^r}(i-x_c)(j-y_c) I_{ij}$,
где $x_c = M_x^r / \sum_{(i,j)\in R^r} I_{ij}$, $y_c = M_y^r / \sum_{(i,j)\in R^r} I_{ij}$ – «центр масс» области R^r.

Шаг 6. Вычислить собственные значения λ_{min}^r и λ_{max}^r и соответствующие собственные векторы v_{min}^r, v_{max}^r для тензора $\mathbf{M}^r = \begin{bmatrix} M_{xx}^r & M_{xy}^r \\ M_{xy}^r & M_{yy}^r \end{bmatrix}$.

Шаг 7. Выполнить: $r := r+1$, $D := D \backslash R^r$.

Шаг 8. Если $r < 3$, то перейти к шагу 3, иначе завершить работу.

Полученные значения характерных размеров (собственные значения) и направлений (собственные векторы) для «тёплых областей» является модельным числовым описанием «длины» и «ширины» областей, являющихся потенциально проблемными участками.

Аналогичная процедура запускается для «холодных» экстремумов, при этом на шаге 3 вместо поиска максимума осуществляется поиск минимума.

Ряд ранее введённых признаков рассчитывается в маске D, полученной в результате «отсечения» «горячих» областей. Это позволяет учесть в модели широкую вариативность форм сосудов у здоровых людей.

6. Отбор признаков

Исходное количество признаков, построенное нами для описания МЖ по термоснимкам, оказалось равным 414, что потребовало проведения дополнительного преобразования признаков и определения их значимости для результатов классификации. Для этого сначала методом главных компонент были построены 200 «главных» признаков, после чего к ним был применен алгоритм последовательного отбора признаков (sequential feature selector), реализованный в библиотеке mlxtend [14]. В так сконструированном векторе признаков в итоге осталось 79 компонент.

7. Решающее правило

В качестве модели для построения решающего правила выбран метод опорных векторов опорных векторов с линейным ядром в виду небольшого объёма обучающей и тестовой выборки. Тестовая выборка содержала 65 образцов (каждый – 3 снимка в разных проекциях), среди которых 12 примеров «ДН», 25 – «ЗН» и 28 – «норма» (без патологий). Модель классификатора была построена и обучена по стратегии «один против всех», а конечные метрики F1-меры,

чувствительности и специфичности были рассчитаны как средние значения по всем бинарным классификаторам.

Результаты и обсуждение

В рамках данной работы поиск оптимального решающего правила на основе метода опорных векторов проходил в 4 этапа.

На первом этапе (I) в качестве базовых признаков были использованы характеристики симметрии, статистические параметры, а также признаки Харалика, рассчитанные на основе анализа матрицы GLCM.

На втором этапе (II) в модель были включены признаки, характеризующие тепловые области изображения, то есть выделялись параметры теплых и холодных областей экстремальных связанных областей.

Далее (III) в модель были добавлены гистограммные признаки, основанные на вычислении расстояния от гистограммы распределения температуры конкретного изображения до усредненных гистограмм для каждого из трёх классов.

На четвёртом этапе (IV) проведён отбор самых информативных признаков, благодаря чему доля правильной классификации (точность) выросла до 86%.

В таблице 1 мы можем проследить динамику увеличения точности классификации в зависимости от используемого набора признаков.

Таблица 1. Результаты поиска оптимального набора признаков

Признаки	Точность	F1-мера	Чувствительность	Специфичность
I	0.72	0.69	0.71	0.75
II	0.75	0.74	0.74	0.71
III	0.79	0.77	0.8	0.75
IV	0.86	0.83	0.85	0.87

Постепенное добавление признаков, связанных с анализом тепловых областей и гистограммных характеристик, существенно улучшило показатели модели по сравнению с базовым вариантом. Это свидетельствует о том, что интеграция термографических данных, отражающих как локальные температурные аномалии, так и глобальное распределение температуры, является эффективным способом повышения качества диагностики. Отбор наиболее значимых признаков позволил дополнительно оптимизировать модель, что снизило риск переобучения и повысило ее надежность при классификации термомаммографических изображений.

Заключение

В данной работе разработан метод автоматизированного анализа термографических изображений молочных желез с применением нейросетевой сегментации и расширенного набора температурных признаков. В отличие от существующих подходов, предложенная методика классифицирует не только наличие новообразования у пациента, но и, предположительно, определяет его природу: доброкачественность или злокачественность. Этим работа отличается от рассмотренных ранее статей [4] и [5], в которых точность классификации выше, однако классификация проходит с целью определения только двух классов: здоров или болен. Выполняя более сложную задачу, мы получили довольно неплохие результаты и достигли точности в 86%, используя при этом значительно меньший набор изображений для обучения, чем авторы вышеуказанных работ. Также, отличительной чертой нашей работы является разработка таких признаков, которые характеризуют температурную асимметрию по квадрантам, корреляцию гистограмм, статистические моменты горячих и холодных областей, а также учитывают исключение сосудистых структур для более точного определения среднетемпературных свойств.

Ограничением исследования является сравнительно небольшой размер выборки, что может влиять на обобщающую способность модели. В дальнейших исследованиях планируется расширение набора данных, внедрение методов глубокого обучения для автоматического извлечения признаков и оптимизация классификационной модели для повышения ее точности и устойчивости.

Литература

1. Всемирная организация здравоохранения. Рак молочной железы [Электронный ресурс]. – 2022. – Режим доступа: https://www.who.int/ru/news-room/fact-sheets/detail/breast-cancer, свободный. – Дата обращения: 10.02.2025.
2. Александрова Л. М., Иванова Т. В., Смирнов П. А. Выявление рака молочной железы: состояние проблемы, пути решения // Онкология. Журнал им. П. А. Герцена. – 2016. – Т. 5, № 2. – С. 34-39.
3. Wang L. Early diagnosis of breast cancer // Sensors. – 2017. – Vol. 17, No. 7. – P. 1572.
4. Banumathy D. et al. CAD of BCD from Thermal Mammogram Images Using Machine Learning // Intelligent Automation & Soft Computing. – 2022. – Vol. 34, No. 1.
5. Periyasamy S. et al. SVM-based classification of thermal images for breast cancer // J. Thermal Biology. – 2022. – Vol. 110. – P. 103337.
6. Prakash J., Khan A. A. Breast Cancer Detection using Thermogram: Review // Int. J. Eng. Res. Technol. – 2016. – Vol. 4, No. 15. – P. 1-12.
7. Ultralytics. Документация (YOLO) [Электронный ресурс]. – Режим доступа: https://docs.ultralytics.com/, свободный. – Дата обращения: 10.02.2025.
8. Redmon J. et al. You Only Look Once: Unified, Real-Time Object Detection // Proc. IEEE CVPR. – 2016. – P. 779-788.
9. Silar. Медицинский тепловизор ТВС300-МЕД [Электронный ресурс]. – Режим доступа: https://www.silar.ru/production/medicinskii-teplovizor-tvs300-med, свободный. – Дата обращения: 03.03.2025.
10. Cao Z. et al. OpenPose: Realtime Multi-Person 2D Pose Estimation Using Part Affinity Fields // arXiv preprint. – 2018.
11. Shih C. L. et al. Difficulty Analyses of Yoga Poses Using OpenPose Keypoints Detection // 2024 Int. Conf. Consumer Electronics-Taiwan (ICCE-Taiwan). – IEEE, 2024. – P. 619-620.
12. Namsin P. et al. Enhanced Image Technique for Yoga Pose Recognition Using MediaPipe and YOLOv8 // 2024 8th Int. Conf. Information Technology (In-CIT). – IEEE, 2024. – P. 174-179.
13. Бухарин Д. Г. и др. Способ прогнозирования риска выявления «малых» форм рака молочной железы: пат. 2547802 РФ. – 2014.
14. Raschka S. mlxtend: Machine Learning Extensions [Электронный ресурс]. – Режим доступа: https://github.com/rasbt/mlxtend, свободный. – Дата обращения: 03.03.2025.

THERMAL IMAGERY MAMMOGRAPHY BCD SCREENING BASED ON COMPUTER VISION AND MACHINE LEARNING

Ph.D. A.V. Khamukhin[1], Prof. S.V. Umnyashkin[1], Prof. I.M. Dolgov[2],
Student D.V. Komoltseva[1], Student A.A. Larev[1]

[1] National Research University of Electronic Technology
[2] LLC "Dignosys"

Classification method proposed for breast thermal images analysis and automatic detection of tumor anomalies with latter being categorized as malignant (cancerous) or benign (noncancerous). The method is based on image segmentation using a neural network model and processing of thermal features in the obtained segments of interest. Over 400 numeric features introduced representing various thermal image descriptors. Evaluation study in this work shows their impact on the quality of patient category classification. For feature importance analysis support vector machine (SVM) is used as backbone method and convenient utility not as prone to overfitting as more complex nonlinear models with huge number of parameters.

The experimental results showed that local thermal anomalies and symmetrical features of temperature distribution in the breasts, with “hot” vascular regions excluded, allows achieving a classification accuracy of 86%, which is fine considering the complexity of 3-category (normal, malignant, benign) classification task. The developed method proved combination of machine learning and thermography to be effective basis for breast cancer detection screening tool.

АВТОМАТИЧЕСКАЯ СЕГМЕНТАЦИЯ МЕТАПЛАЗИИ В ЭНДОСКОПИЧЕСКОЙ СИСТЕМЕ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ

проф. Обухова Н.А. [1,2], доц. Мотыко А.А.[1,2], асп. Савельев А.Г.[1], асп. Савельева О.Р.[2], доц. Самарин А.В. [3], зав. отд. эндоскопии, к.м.н. Сигаева Ю.С.[4]

[1] Санкт-Петербургский государственный электротехнический университет «ЛЭТИ»
[2] Санкт-Петербургский государственный университет
[3] Национальный исследовательский университет ИТМО, г. Санкт-Петербург
[4] «ММЦ», г. Геленджик

Целью исследования является разработка нейросетевой модели для автоматической сегментации изображений, в рамках реализации системной функции поддержки принятия врачебных решений в видеоэндоскопической системе. В основе разработки лежит использование современных технологий искусственного интеллекта, а именно нейронных сетей, которые обеспечивают автоматическую сегментацию метаплазии на эндоскопических снимках. Метаплазия, являясь предраковым состоянием, требует особого внимания со стороны врачей, и наличие инструмента, способного поддерживать их при проведении обследования, имеет важное значение для улучшения диагностики и повышения уровня медицинской помощи. В рамках исследования была собрана база данных изображений, предложена оригинальная архитектура сети и создана нейросетевая модель, которая позволила добиться целевых результатов в сегментации метаплазии. Проведенные исследования показали, что разработанная модель превосходит по точностным характеристикам современные аналоги (на 0,04-0,11 по метрике IoU). Использование результатов исследования позволяет ускорить процесс анализа данных при проведении видеоэндоскопического обследования, способствует снижению вероятности врачебной ошибки, что является важным шагом к улучшению клинических результатов для пациентов.

Введение

В последние годы эндоскопия играет ведущую роль в диагностике и лечении заболеваний внутренних органов, особенно в области гастроэнтерологии. Эффективность эндоскопических исследований во многом зависит от точности интерпретации полученных изображений, что требует от врачей высокой квалификации и значительных временных затрат. В связи с ростом объема эндоскопических данных и увеличением нагрузки на медицинский персонал, возникает необходимость в разработке высокотехнологичных устройств, способных оказывать информационную поддержку врачу при проведении обследовании и принятии решений.

Эндоскопическое исследование желудочно-кишечного тракта (ЖКТ) играет ключевую роль в профилактике заболеваний, таких как рак желудка. Рак желудка относится к онкологическим заболеваниям с неблагоприятным прогнозом и часто остается незамеченным при стандартных осмотрах. Современные методы диагностики не всегда позволяют выявить предраковые состояния и ранние стадии рака, что ограничивает возможности лечения и снижает показатели выживаемости пациентов. Согласно данным Международного агентства по изучению рака (МАИР) [1], специализированного агентства Всемирной организации здравоохранения (ВОЗ), рак желудка остается серьезной глобальной проблемой для общественного здравоохранения. Каждый год регистрируется 968 784 новых случаев рака желудка и 660 175 смертей, связанных с этим заболеванием, по всему миру [2]. Гастроскопия используется как скрининговая процедура для диагностики заболеваний верхних отделов ЖКТ, включая рак, особенно в регионах с высоким риском. Основная цель скрининга с помощью гастроскопии — повышение выявляемости рака желудка на ранних стадиях и снижение смертности от этого заболевания. Однако успех этих задач во многом зависит от опыта врача-эндоскописта. Исследования показывают, что 20–25% ранних стадий рака желудка могут быть пропущены, а 11,3% случаев рака верхних отделов ЖКТ на поздних стадиях остаются невыявленными. Востребованность скрининга и существующая острота проблемы обуславливает актуальность разработки программных решений, направленных на оказание врачу поддержки при проведении обследования.

В данной работе обращено внимание на желудочно-кишечную метаплазию, которая считается предраковым состоянием. При метаплазии слизистая оболочка желудка, вырабатывающая соляную кислоту, трансформируется и становится похожей на слизистую кишечника. Выявление метаплазии затруднено из-за ее плоской морфологии и незначительных отличий от нормальной слизистой. Автоматическому обнаружению и сегментации метаплазии посвящены следующие

публикации: в работе [3] описано решение задачи сегментации желудочно-кишечной метаплазии, разработанная система выполняет анализ в реальном времени для нескольких режимов визуализации с помощью SAM (Segment Anything Model); статья [4] демонстрирует работу нескольких архитектур сверточной нейронной сети (ResNet50, VGG 16, VGG 19 и Inception V3) для обнаружения кишечной метаплазии; в [5] исследование посвящено распознаванию атрофического гастрита и кишечной метаплазии на эндоскопических изображениях в белом свете (WLI) на основе сверточной нейронной сети TResNet.

В данной работе представлено оригинальное исследование, посвященное сегментации желудочно-кишечной метаплазии. Исследование включало сбор и разметку целевой базы данных. При выполнении эндоскопического обследования метаплазия плохо различима в стандартном режиме – в белом свете. Для ее обнаружения врач обычно использует специальные режимы визуализации, в которых значительно повышен локальный цветовой контраст и существенно изменена цветовая палитра. Проблемы заключаются в следующем: многие эндоскопы не обладает возможностью такой визуализации; в ряде случаев врач может пропустить малоразмерный очаг метаплазии там, где он его не ожидает и, соответственно, не активирует специальный режим. Последняя проблема усугубляется в случае перегруженности врача, его усталости, недостаточной опытности и т.д.

Данное исследование направлено на создание решения, применимого на практике в широкой линейке оборудования, без использования специализированной визуализации.

Предложенное решение

В данной работе предлагается нейросетевая архитектура для сегментации метаплазии, основанная на модифицированной модели U-Net [6], работающей с цветными трехканальными изображениями. Идея архитектуры заключается в объединении высокой эффективности U-Net в задачах сегментации и возможностями блоков усиленного самовнимания, предложенными в работе SSD-ADSAR [7], а также включением дополнительной ветви, осуществляющей классификацию, предложенной в [8]. Кодирующая часть (L-Net) U-Net извлекает базовые признаки изображения, после чего они дополняются результатом работы модуля самовнимания, позволяющим учитывать как локальные детали, так и глобальный контекст изображения. Полученное представление используется для формирования двух выходов – сегментационной маски, локализующей области метаплазии, и бинарного предсказания наличия патологических изменений. При обучении применяется комбинированная функция потерь, объединяющая термы, отвечающие за сегментацию, классификацию и регуляризацию, что способствует повышению устойчивости модели к вариациям во входных данных. Представленная в данной работе нейросеть была названа авторами U-Net ADSAR.

Модуль самовнимания внедрён в кодирующую часть U-Net и позволяет усилить выделение признаков за счёт разделения внимания на локальный и глобальный потоки. Пусть входное представление $\mathbf{X} \in \mathrm{R}^{H\times W\times D}$ представляет собой набор признаков, полученных после начальных сверточных операций. Для дальнейшей обработки признаки преобразуются в запросы **Q**, ключи **K** и значения **V** посредством обучаемых линейных преобразований:

$$\boldsymbol{Q} = \boldsymbol{X}\boldsymbol{W}^{\boldsymbol{Q}}, \boldsymbol{K} = \boldsymbol{X}\boldsymbol{W}^{\boldsymbol{K}}, \boldsymbol{V} = \boldsymbol{X}\boldsymbol{W}^{\boldsymbol{V}}, \quad (1)$$

где $\mathbf{W}^{\mathbf{Q}}, \mathbf{W}^{\mathbf{K}}, \mathbf{W}^{\mathbf{V}} \in \mathrm{R}^{D\times D'}$ — матрицы преобразования.

Для разделения внимания на локальный и глобальный потоки вводятся две маски. Локальная маска $\mathbf{M}_\mathbf{l}$ активирует признаки, находящиеся в пределах фиксированной окрестности *N(i)* для каждой позиции *i*, а глобальная маска $\mathbf{M}_\mathbf{g}$, наоборот, подавляет информацию из этой области:

$$M_l(i,j) = \begin{cases} 0, \text{если } j \in N(i), \\ -\infty, \text{ иначе,} \end{cases} \quad M_g(i,j) = \begin{cases} 0, \text{если } j \notin N(i), \\ -\infty, \text{ если } j \in N(i). \end{cases} \quad (2)$$

Веса для локального и глобального потоков рассчитываются с использованием функции softmax:

$$A_l = softmax\left(\frac{QK^T}{\sqrt{D'}} + M_l\right), \qquad A_g = softmax\left(\frac{QK^T}{\sqrt{D'}} + M_g\right) \tag{3}$$

Полученные отклики вычисляются следующим образом:

$$R_l = A_l V, R_g = A_g V. \tag{4}$$

Особенностью предлагаемого подхода является использование абсолютной разности между локальными и глобальными потоками для усиления релевантных признаков:

$$Res = \left|R_l - R_g\right|, \tag{5}$$

после чего результирующее представление интегрируется с исходными признаками:

$$X' = X + Res \cdot \mathbf{F}, \tag{6}$$

где $\mathbf{F} \in R^{D' \times D}$ — обучаемая матрица проекции, возвращающая представление в исходное размерное пространство.

Идея внедрения такого модуля была предложена в работе [7], где доказано, что разделение внимания на локальный и глобальный потоки позволяет существенно улучшить детальное представление признаков.

Дополнительно в архитектуру встроена классификационная ветвь, реализованная в виде многоуровневого перцептрона, которая принимает сжатое представление из L-Net и обучается предсказывать наличие патологических изменений. Этот компонент был разработан с учётом подхода, описанного в работе [8], что позволило добиться высоких результатов при объединении задач сегментации и классификации.

Для повышения устойчивости модели к вариативности входных данных в функцию потерь добавлена так называемая регуляризация согласованности (Consistency Regularization [9]). Итоговая функция потерь имеет вид:

$$L_{total} = L_{seg} + L_{cls} + \alpha \cdot f\big(I, aug(I)\big), \tag{7}$$

где L_{seg} терм, отвечающий за точность сегментации, L_{cls} терм, отвечающий за точность классификации, α — коэффициент регуляризации, а функция *f(I, aug(I))* измеряет расхождение между предсказаниями для исходного изображения *I* и его аугментированной версии *aug(I)*.

При этом после развертывания используется только кодирующая часть сети с классификационной ветвью, что позволяет снизить вычислительную сложность без потери качества. На рисунке 1 изображена предложенная архитектура.

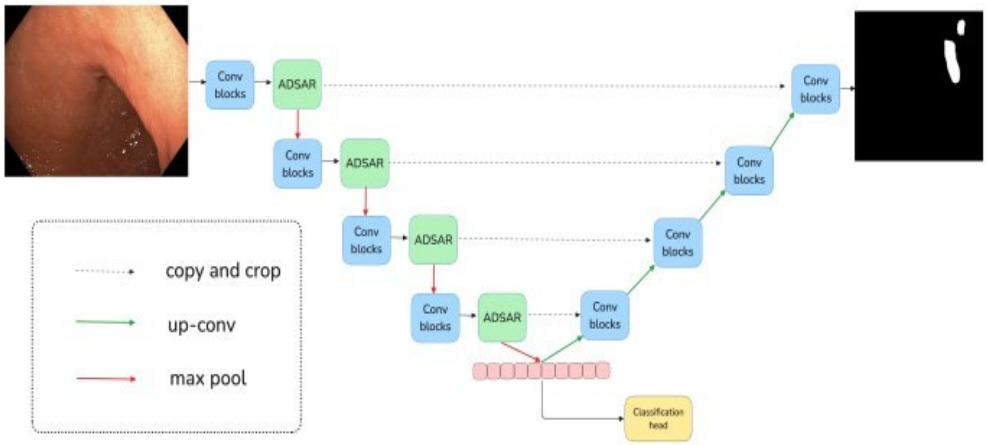


Рисунок 1 – Схема предлагаемой архитектуры U-Net ADSAR

Набор данных

В рамках исследования эндоскопические данные были собраны и аннотированы опытным эндоскопистом. В общей сложности было собрано около 2300 эндоскопических снимков. Одной из основных проблем в ходе сбора данных было неодинаковое визуальное качество получаемых изображений, обусловленное условиями их формирования в ходе реальных медицинских

процедур и операций. Это повлекло за собой необходимость отбраковки около 300 изображений, и соответствующего сокращения объема базы данных. На каждом изображении полученной выборки было выделено от 3 до 5 объектов интереса. Результаты разметки представляют собой маски, соответствующие очагам метаплазии. Для обеспечения высокого качества снимков были разработаны специальные методические рекомендации для врача и разметчиков, что позволило систематизировать процесс сбора данных и минимизировать ошибки в аннотациях. Выборка была разделена на обучающий (1600 изображений) и тестовый (400 изображений) наборы.

Примеры изображений и разметки из собранной базы данных приведены на рисунке 2.

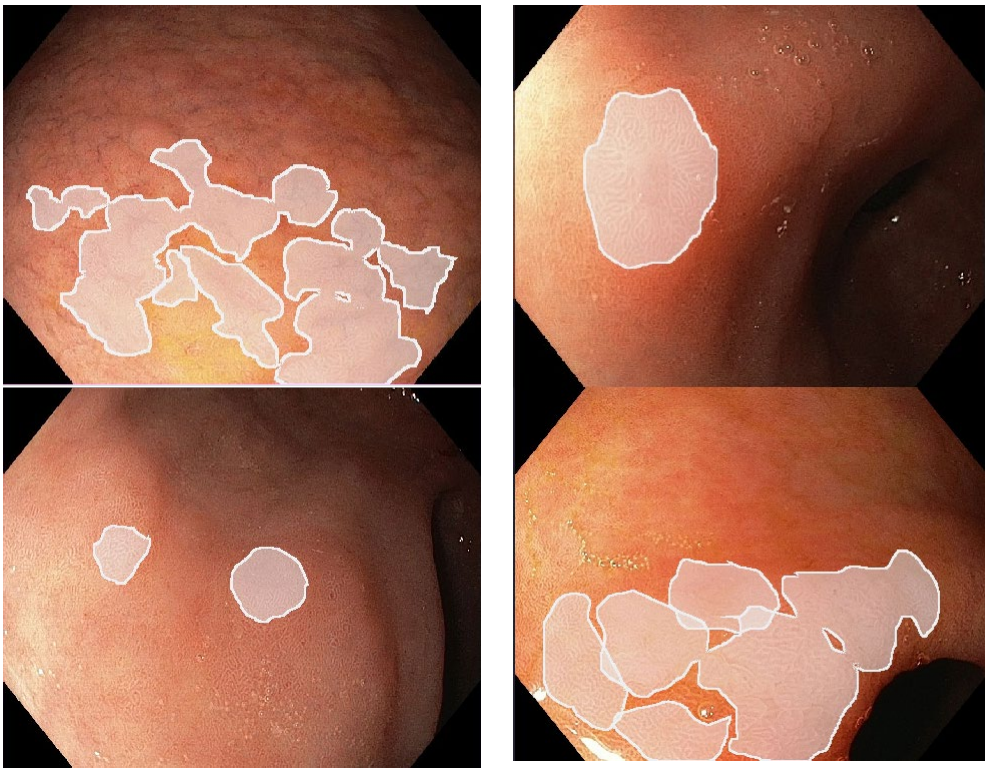


Рисунок 2 - Примеры изображений из базы данных

Базовые модели

Для объективной оценки эффективности предложенной архитектуры U-Net ADSAR и вклада техники регуляризации согласованности нами были обучены модели, представляющие различные подходы к сегментации метаплазии на RGB-изображениях. Такой разнообразный выбор базовых решений позволяет сравнить результаты с классическими архитектурами, а также с легковесными моделями, ориентированными на быструю классификацию патологических изменений.

В первую очередь, в экспериментах использовались стандартные модели U-Net с различными базовыми энкодерами. Так, модель U-Net ResNet-50 была выбрана за её доказанную эффективность в ряде медицинских задач, где базовая сеть ResNet-50 обеспечивает баланс между точностью и вычислительной сложностью. Аналогично, модель U-Net EfficientNetB4 была включена для демонстрации преимуществ современных архитектур EfficientNet, способных извлекать более сложные признаки при меньшем количестве параметров.

Для более глубокого анализа влияния техники регуляризации согласованности нами были обучены модифицированные версии указанных моделей, где в функцию потерь было добавлено слагаемое, минимизирующее различия между предсказаниями для исходного изображения и его аугментированной версии. Таким образом, получены модели U-Net ResNet-50 с consistency loss и U-Net EfficientNetB4 с consistency loss, позволяющие оценить вклад данного подхода в повышение устойчивости и обобщающей способности сети.

В качестве дополнительного базового решения был использован подход, предложенный в работе [8]. Из данной работы была выделена компактная модель LNet, ориентированная на эффективное извлечение признаков и классификацию патологических изменений, что позволяет оценить качество представлений в условиях ограниченных вычислительных ресурсов.

Основное внимание в исследовании было уделено предлагаемым решениям, разработанным на базе U-Net ADSAR. Предлагаемая архитектура UNet ADSAR представляет собой

модифицированную версию U-Net, в кодирующую часть которой интегрирован модуль самовнимания, что позволяет эффективно выделять как локальные, так и глобальные признаки, характерные для метаплазии. Кроме того, разработана версия U-Net ADSAR с consistency loss, где наряду с модулем самовнимания в функцию потерь включена регуляризация согласованности для повышения устойчивости модели к вариациям входных данных.

Детали обучения

Все модели были обучены на GPU NVIDIA RTX 3090 до 100 эпох. Оптимизация выполнялась с использованием оптимизатора AdamW [10] с параметрами $\beta1 = 0.9$ и $\beta2 = 0.999$, и затуханием весов 0.05. Размер батча был установлен равным 16. Начальная скорость обучения была установлена на 10^{-8}, с использованием косинусного графика уменьшения скорости обучения до минимума 0 по мере продвижения обучения. Кроме того, каждая модель прошла 10 повторений обучения, чтобы повысить надежность и стабильность выводов, сделанных во время тестовой фазы. Для обучения модели U-Net ADSAR использовалось 4 головы в многоуровневом двупоточном самовнимании.

Экспериментальные результаты и выводы

Для оценки производительности моделей в задаче сегментации метаплазии использовалась метрика пересечение над объединением (IoU). IoU является популярной метрикой, измеряющей степень совпадения между предсказанной и истинной масками, рассчитываемой как отношение площади их пересечения к площади их объединения:

$$IoU = \frac{|X \cap Y|}{|X \cup Y|}, \tag{8}$$

где X — это предсказанный результат сегментации, Y — это истинный результат сегментации.

Таблица 1 содержит сводные результаты по метрике IoU для каждого из тестируемых решений.

Таблица 1. Сравнение моделей сегментации метаплазии

Метод	**IoU**
U-Net ResNet-50	0.77
U-Net EfficientNetB4	0.80
U-Net ResNet-50 с consistency loss	0.82
U-Net EfficientNetB4 с consistency loss	0.84
U-Net [8]	0.84
U-Net ADSAR	0.86
U-Net ADSAR с consistency loss	0.88

Анализ полученных результатов показывает, что модели, использующие предложенную модификацию U-Net ADSAR, демонстрируют лучшие точностные показатели. U-Net ADSAR превосходит аналоги на 0,02-0,09 по метрике IoU, а версия с consistency loss на 0,04-0,11 по метрике IoU. Следует отметить, что добавление модуля усиленного самовнимания в кодирующую часть существенно улучшает точность локализации патологических областей, а применение регуляризации согласованности положительно сказывается на устойчивости модели к вариациям входных данных. Наиболее высокие значения IoU были достигнуты при использовании U-Net ADSAR с consistency loss, что подтверждает эффективность интеграции данных компонент в архитектуру.

Таким образом, в результате исследования была создана оригинальная нейросетевая архитектура и получена соответствующая модель для автоматической сегментации изображений, в рамках реализации системной функции поддержки принятия врачебных решений в видеоэндоскопической системе. Результаты экспериментов свидетельствуют о значительном улучшении точности сегментации метаплазии с помощью разработанной модели по сравнению с аналогами, что достигается за счёт объединения возможностей классических архитектур U-Net с современными подходами к усилению представлений.

Источник финансирования

Исследование (раздел “экспериментальные результаты и выводы”) выполнено при финансовой поддержке Министерства науки и высшего образования Российской Федерации в рамках реализации научного проекта по соглашению № 075-15-2025-013

Литература

1. Cancer Today [Electronic resource]. URL: https://gco.iarc.who.int/today/en (дата обращения: 06.02.2025).
2. Bravo D. et al. GastroHUN an Endoscopy Dataset of Complete Systematic Screening Protocol for the Stomach // Scientific Data 2025 12:1. Nature Publishing Group, 2025. Vol. 12, № 1. P. 1–14.
3. Pornvoraphat P. et al. Real-time gastric intestinal metaplasia segmentation using a deep neural network designed for multiple imaging modes on high-resolution images // Knowl Based Syst. Elsevier, 2024. Vol. 300. P. 112213.
4. Ligato I. et al. Convolutional Neural Network Model for Intestinal Metaplasia Recognition in Gastric Corpus Using Endoscopic Image Patches // Diagnostics 2024, Vol. 14, Page 1376. Multidisciplinary Digital Publishing Institute, 2024. Vol. 14, № 13. P. 1376.
5. Lin N. et al. Simultaneous Recognition of Atrophic Gastritis and Intestinal Metaplasia on White Light Endoscopic Images Based on Convolutional Neural Networks: A Multicenter Study // Clin Transl Gastroenterol. Clin Transl Gastroenterol, 2021. Vol. 12, № 8. P. E00385.
6. Ronneberger O., Fischer P., Brox T. U-net: Convolutional networks for biomedical image segmentation // Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). Springer Verlag, 2015. Vol. 9351. P. 234–241.
7. Samarin A. et al. ADSAR: Advanced Dual-Stream Attention and Reweighting for Small Object Detection.
8. Samarin A. et al. One-Staged Attention-Based Neoplasms Recognition Method for Single-Channel Monochrome Computer Tomography Snapshots // Pattern Recognition and Image Analysis. Pleiades journals, 2022. Vol. 32, № 3. P. 645–650.
9. Samarin A. et al. Predictors Based on Convolutional Neural Networks for the Movement Strategy of Trainable Agents for Building Customized Image Descriptors // Pattern Recognition and Image Analysis. Pleiades Publishing, 2023. Vol. 33, № 2. P. 139–146.
10. Loshchilov I., Hutter F. Decoupled Weight Decay Regularization // 7th International Conference on Learning Representations, ICLR 2019. International Conference on Learning Representations, ICLR, 2017.

AUTOMATIC METAPLASIA SEGMENTATION IN AN ENDOSCOPIC DECISION SUPPORT SYSTEM

N. Obukhova [1,2], A. Motyko[1,2], A. Savelev[1], O. Saveleva[2],
A. Samarin [3], J. Sigaeva[4]

[1] Saint Petersburg Electrotechnical University "LETI"
[2] Saint Petersburg State University
[3] ITMO University, Saint Petersburg
[4]Group of companies “My Medical Center”, Gelendzhik

The aim of the study is to develop a neural network model for automatic segmentation of images as a part of the implementation of medical decision support function in video endoscopy system. The development is based on the use of modern artificial intelligence technologies, namely neural networks that provide automatic segmentation of metaplasia on endoscopic images. Metaplasia, being a precancerous condition, requires special attention from physicians, and having a tool that can support them in their examinations is essential to improve diagnosis and enhance care. In this study, an image database was collected, an original network architecture was proposed and a neural network model was developed to achieve targeted results in metaplasia segmentation. The conducted research showed that the developed model outperforms the state-of-the-art competitors in terms of accuracy (by 0.04-0.11 on the IoU metric). Using the results of the study allows to speed up the process of data analysis during video endoscopic examination, contributes to reducing the probability of medical error, which is an important step towards improving clinical outcomes for patients.

ПОВЫШЕНИЕ ЭФФЕКТИВНОСТИ ОБНАРУЖЕНИЯ БПЛА НА ОСНОВЕ РАДИОЛОКАЦИОННЫХ ДАННЫХ С ИСПОЛЬЗОВАНИЕМ НЕЙРОСЕТЕВЫХ МЕТОДОВ

к.т.н. Балакин Д.А.[1], инж. Семичастнов А.Е.[1], инж. Горбунов А.Р.[1], инж. Акмалов М.Р.[1]

[1] ЦИТМ «ЭКСПОНЕНТА»

Работа посвящена повышению эффективности обнаружения беспилотных летательных аппаратов (БПЛА) на основе радиолокационных данных с использованием нейросетевых методов. Актуальность исследования обусловлена проблемой ложных срабатываний в системах обнаружения из-за схожести размеров БПЛА и птиц. Для решения задачи предложен метод классификации, основанный на анализе микродоплеровских сигнатур (МДС), которые отражают динамику движения объектов. Моделирование радиолокационных сигналов выполнено в среде Engee, где разработаны цифровые модели БПЛА и птиц. Сформирован синтетический набор данных из нескольких тысяч изображений для каждого класса, включающий спектрограммы МДС. Для классификации были использованы архитектуры EfficientNet (B0 и B7), а также метод дистилляции знаний. Исследование демонстрирует потенциал интеграции нейросетевых алгоритмов в системы защиты критически важных объектов, обеспечивая минимизацию ложных срабатываний. Перспективы работы связаны с расширением базы данных за счет реальных сигналов и интеграцией комплексных методов обнаружения.

Введение

Беспилотные летательные аппараты (БПЛА) находят широкое применение как в гражданской, так и в военной сферах, что делает актуальной задачу их надежного обнаружения для защиты критически важных объектов [1]. Одним из наиболее эффективных методов обнаружения БПЛА является радиолокация, которая позволяет осуществлять мониторинг на больших расстояниях и в любых погодных условиях. Однако схожесть размеров БПЛА и птиц часто приводит к ложным срабатываниям, что создает серьезные трудности. Данная проблема требует разработки новых подходов, способных повысить точность обнаружения и минимизировать количество ошибок [2].

Для обнаружения БПЛА применяются различные методы: радиочастотный анализ, акустические датчики, оптические системы и анализ радиочастотного тракта. Анализ радиочастотного тракта (например, идентификация сигналов управления) теряет актуальность с развитием автономных БПЛА. Акустические системы чувствительны к шумам окружающей среды, а оптические камеры ограничены дальностью действия и ухудшением работы в плохих погодных условиях. Комплексный подход, объединяющий данные от нескольких источников, позволяет повысить надежность обнаружения за счет взаимной компенсации недостатков отдельных методов. Однако ключевую роль в таких системах играют радиолокационные технологии, обеспечивающие всепогодность и дальность обнаружения [3].

Радиолокационный подход

В работе обнаружение БПЛА и птиц осуществляется на основе микродоплеровских сигнатур. На основе анализа радиолокационного подхода разработана схема (рисунок 1) обработки принятого сигнала, которая учитывает ограниченность объема реальных данных [3].

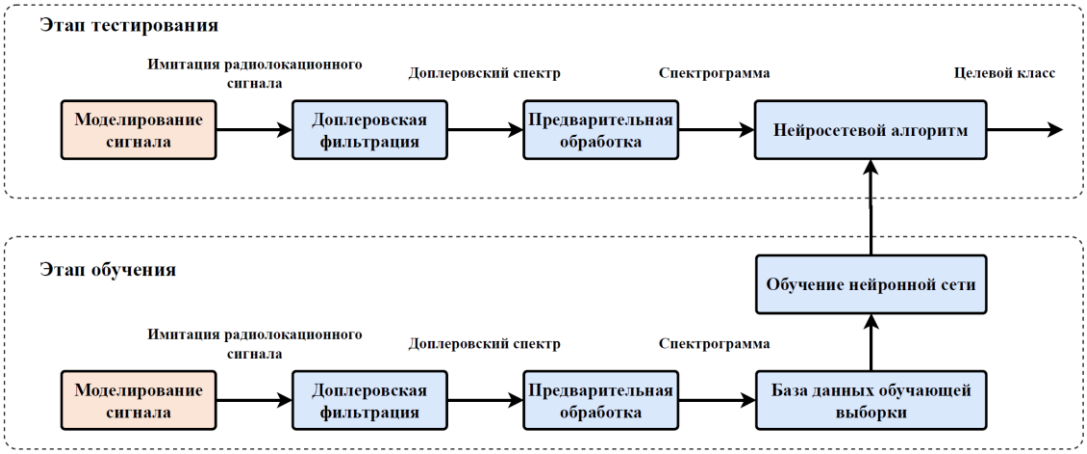


Рисунок 1 – Преобразованная схема обработки радиолокационного сигнала

В связи с этим набор данных для обучения нейронной сети формируется с использованием синтетического генератора, основанного на смоделированных сигналах. Сбор реальных данных о микродоплеровских сигнатурах БПЛА и птиц затруднен из-за ограниченного доступа к различным типам беспилотников и необходимости проведения сложных натурных экспериментов. Кроме того, реальные данные часто охватывают узкий спектр условий, таких как конкретные типы БПЛА, погодные параметры или расстояния, что снижает их применимость для обучения универсальных моделей.

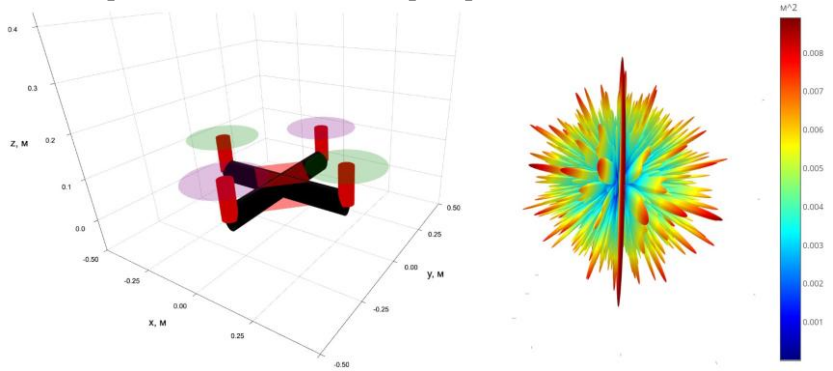
Таким образом, синтетические данные, генерируемые в среде моделирования Engee, становятся ключевым инструментом. Детали формирования модельных данных будут подробно рассмотрены в следующем разделе.

Формирование модельных сигналов

В данной работе выполнен анализ наиболее распространённых типов беспилотных летательных аппаратов, включая вертолетные и мультироторные. Данные типы БПЛА часто используются для разведывательных задач благодаря своей высокой маневренности и способности оставаться незамеченными.

Для формирования диаграммы обратного рассеяния сложной цели необходимо разбить анализируемый объект на ключевые компоненты, которые учитывают микродоплеровский эффект. В результате моделирования в среде Engee были получены цифровые модели, диаграммы обратного рассеяния (ДОР), осциллограммы и спектрограммы сигналов [4,5].

Для представления БПЛА типа квадрокоптер в упрощенной форме необходимо разделить на следующие основные компоненты: четыре двухлопастных ротора (заменены на 2 плоские пластины) и корпус БПЛА, представленный цилиндрическими структурами. Геометрия цифровой модели показана на рисунок 2а. В результате выполнения моделируемого сценария, где объект облучался со всех сторон, диаграмма обратного рассеяния принимает вид, показанный на рисунок 2б. Из рисунка видно, что структура ДОР образует иглообразную форму в направлении роторов БПЛА. Эта характеристика объясняется интерференционным эффектом, возникающим из-за отражений от нескольких роторов БПЛА.



а) б)

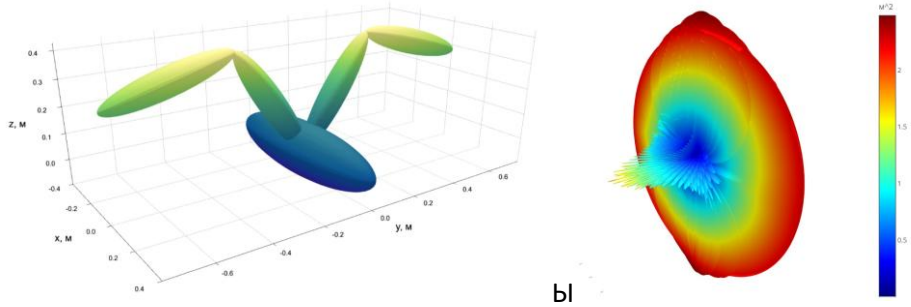
Рисунок 2 –Модель (а), диаграмма обратного рассеивания (б) БПЛА типа квадрокоптер

При описании птицы как объекта необходимо учитывать орнитологические параметры и динамику ее движения. Структура крыла определяется следующими элементами: плечевая кость (верхняя часть крыла), предплечье (локтевая и лучевая части) и кисть [6]. Взмахи крыльев характеризуются подъемом и опусканием верхней части предплечья вокруг суставов. Ключевыми элементами крыла, формирующими микродопплеровские сигнатуры, являются плечевая кость и предплечье. Кисть практически не вносит вклад в спектр отраженного сигнала из-за своих малых размеров, поэтому для упрощения предплечье и кисть объединены в один элемент. На основе этих соображений был сформирован внешний вид приблизительной цифровой модели птицы (рисунок 3а). Наиболее точное приближение компонентов птицы достигается с использованием эллиптических поверхностей, для которых существует аналитическое выражение эффективной площади рассеяния (ЭПР), заданное формулой (1).

$$\sigma_{\text{эллипс}} = \frac{\pi a^2 b^2 c^2}{a^2 \sin(\theta)^2 + b^2 \sin(\theta)^2 + c^2 \cos(\theta)^2} \quad (1)$$

где a, b, c – длины полуосей эллипсоида, θ – полярный угол.

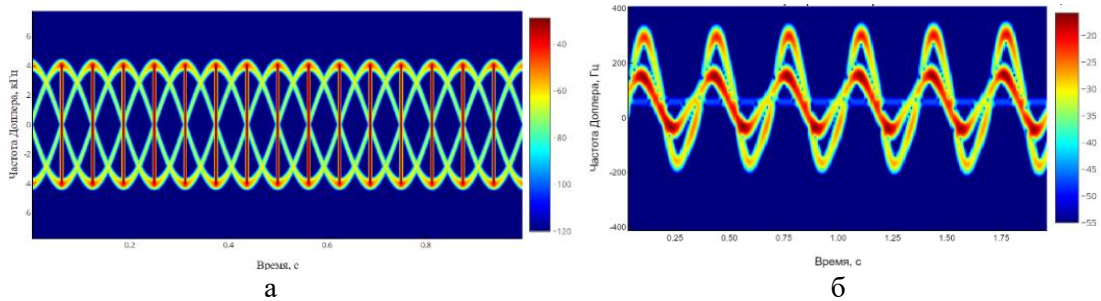
На рисунке 3б изображена диаграмма обратного рассеивания модели птицы.



а) б)

Рисунок 3 – Модель (а), диаграмма обратного рассеивания (б) птицы

Используя выражения, взятые из [7] построим спектрограмму вертолета с четырьмя винтами при частоте вращения $f_{\text{вращ}}$ = 4 Гц приведен на рисунке 4а. Также используя выражения из [7] построим спектрограммы отраженного сигнала от птицы частотой взмаха крыльев $f_{\text{взмах}}$ = 3 Гц, приведем полученные результаты на рисунке 4б.

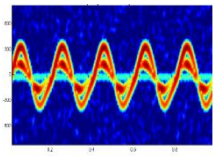
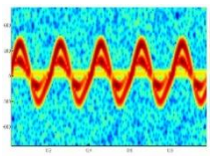
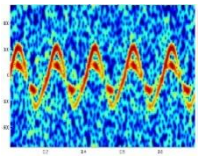
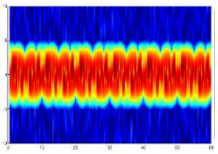
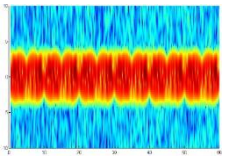
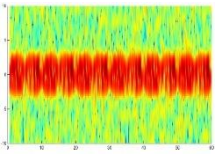
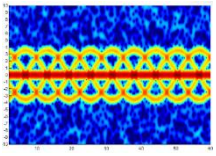
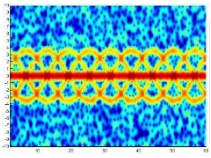
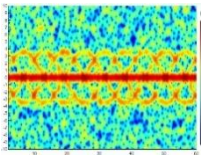
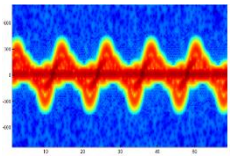
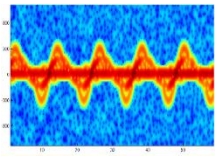
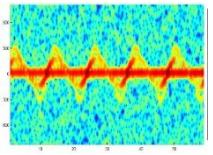


а б

Рисунок 4 – Спектрограмма отраженного сигнала: а) БПЛА вертолетного типа; б) птицы

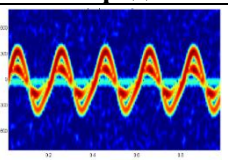
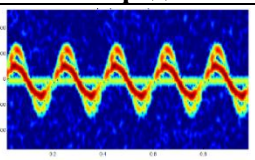
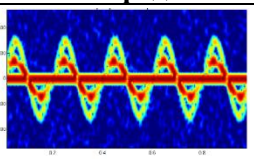
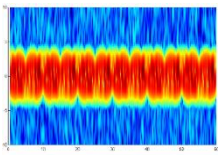
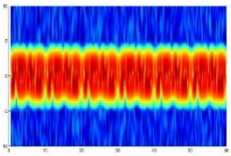
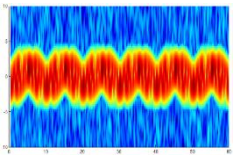
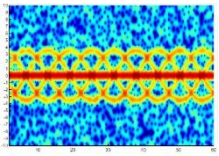
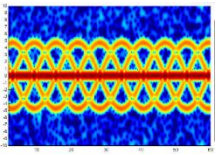
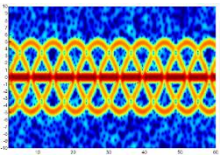
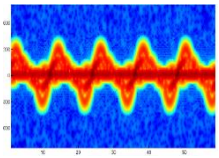
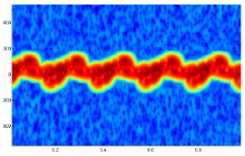
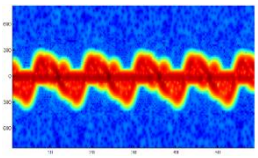
В таблицах сведены наиболее наглядные результаты работы различных сценариев движения объекта. В таблице 1 сведем спектрограммы при варьировании дальности до объекта.

Таблица 1. Спектрограммы при варьировании дальности до объекта

Тип объекта	**Дальность до объекта**		
	200 м	**400 м**	**800 м**
Птица			
БПЛА типа квадрокоптер			
БПЛА вертолетного типа			
БПЛА самолетного типа			

Анализируя таблицу1, можно сделать вывод, что увеличение дальности приводит к большему влияние шума приемника. В таблице 2 сведем спектрограммы при варьировании направления движения объекта относительно движения РЛС.

Таблица 2. Спектрограммы при варьировании направления движения объекта

Тип объекта	Направление движения объекта относительно РЛС		
	0 град.	45 град.	90 град.
Птица			
БПЛА типа квадрокоптер			
БПЛА вертолетного типа			
БПЛА самолетного типа			

Путем варьирования параметров таких как: дальность объекта, направление движения объекта относительно РЛС, скорость объекта, скоростью вращения винтов, частота взмахов крыльев, условий и среды распространения было получено несколько тысяч изображений для каждого типа БПЛА, что позволило сформировать необходимые данные для последующего шага –повышение эффективности обнаружения БПЛА с помощью нейросетевого алгоритма.

Глубокое обучение для классификации спектрограмм

В ходе исследования спектрограммы, представленные в таблицах 1 и 2, интерпретировались как набор изображений, что позволило использовать аппарат нейронных сетей (НС).

Одна из наиболее популярных архитектур НС – EfficientNet [8], имеющая две версии: B0 и B7. EfficientNet-B0 имеет относительно небольшое число варьируемых параметров (около 5,3M), позволяющая быстро обучаться, но с невысокой точностью. EfficientNet-B7, напротив, содержит около 66M параметров и имеет более высокое качество работы, что сказывается на быстродействии. Рассмотрев положительные и отрицательные стороны B0 и B7, было принято решение применить подход дистилляции знаний, который позволяет дообучить менее точную НС – B0 («Student») с помощью более точной B7 («Teacher»). Процесс дистилляции показан на рисунке 5, используя предсказания «Teacher», «Student» корректировать свои прогнозы.

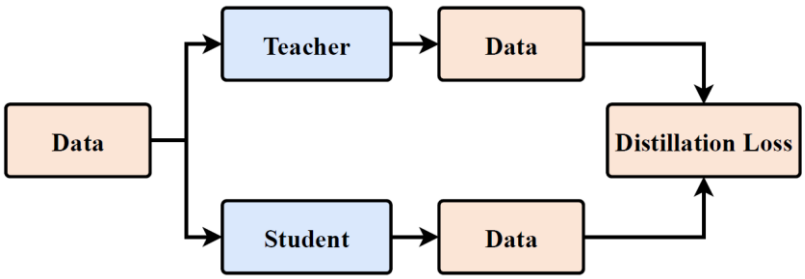


Рисунок 5 – Структурная схема процесса дистилляции

В результате «Student» улавливает взаимосвязи между классами, улучшая обобщение модели. Soft targets модели-учителя получаются с использованием функции softmax с температурой [9].

$$P_i^T = softmax\left(\frac{Z_i}{T}\right) \quad (2)$$

где Z_i – «logits» модели-учителя, T – параметр температуры, используемый для смягчения распределения вероятностей.

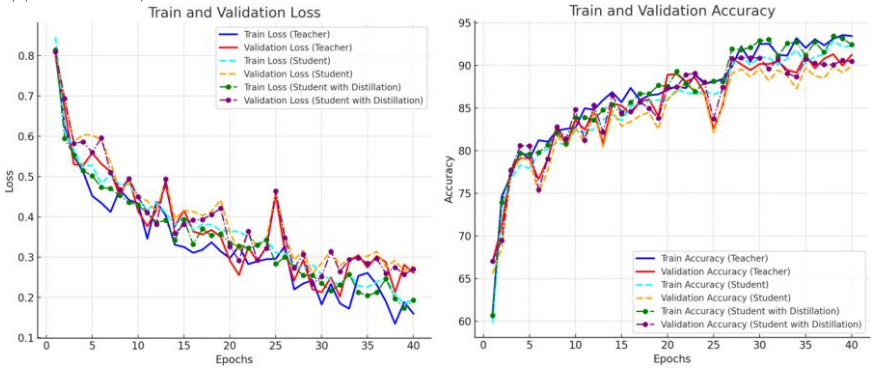
«Student» обучается с функцией потерь, комбинирующей кросс-энтропию и дивергенцию Кульбака-Лейблера между soft targets «Teacher» и предсказаниями «Student», что улучшает точность модели. Дистилляционная функция потерь в соответствии с [9] имеет вид:

$$L = \alpha \cdot L_{CE}(y, y_{student}) + (1 - \alpha)T^2 KL \quad (3)$$

где L_{CE} – кросс-энтропия, KL – дивергенция Кульбака-Лейбнера, α – гиперпараметр, регулирующий вклад отдельного компонента L, T – температура.

Результаты

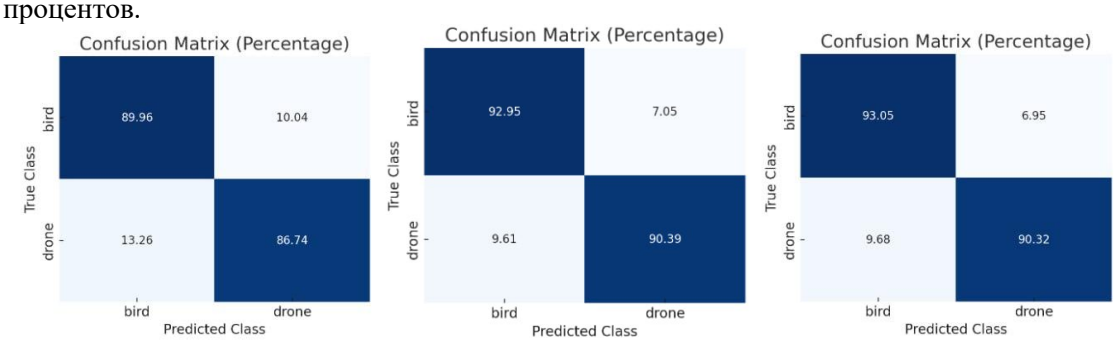
Модели EfficientNet-B0 и EfficientNet-B7 и их комбинации с дистилляцией знаний были обучены на синтетических данных. Результаты представлены на рисунке 6, где описываются точность модели и ошибка на тренировке и валидации. Ошибка «Teacher» достигает более низких значений по сравнению с другими. «Student» с дистилляцией показывает меньшую ошибку как на обучении, так и на валидации по сравнению с B0, что подтверждает эффективность дистилляции.



а б

Рисунок 6 – График ошибки (а), точности (б) на обучающей и валидационной выборках

Ниже (рисунок 7) представлены результаты классификации объединенного класса «Дрон» (объединяющий БПЛА из таблицы 1 и 2) по сравнению с классом «Птица», где точность выше 90 процентов.



а б в

Рисунок 7 – Результат работы: а) студент; б) студент после дистилляции; в) учитель.

Для оценки эффективности обучения модели использовалась метрика Accuracy. Модель «Student» без применения дистилляции показала точность на валидационной выборке 88.5%, что несколько ниже результата модели «Teacher» (91.4%). После применения дистиляции точнсоть модели «Student» увеличилась до 91.6% , незначительно превысив показатель модели «Teacher».

Заключение

В работе представлен подход к повышению эффективности обнаружения БПЛА от птиц на основе радиолокационных данных с использованием нейросетевого подхода. В среде моделирования Engee были разработаны цифровые модели различных типов БПЛА (вертолетных, мультироторных) и птиц. Данные цифровые модели позволили преодолеть ограничения, связанные с недостатками реальных данных и сформировать базу данных для обучения нейронной сети.

Представленный алгоритм, основанный на глубоком обучении, демонстрирует высокую точность (более 90%) в обнаружении БПЛА и птиц. В качестве классификаторов применялись модели EfficientNet (B0 и B7), которые показали высокую эффективность в обработке спектрограмм. Также был рассмотрен метод дистилляции знаний, позволяющий улучшить точность более простых моделей за счет обучения на предсказаниях более сложных. Результаты работы демонстрируют, что предложенный подход может быть успешно применен для создания систем обнаружения и защиты от БПЛА, особенно в условиях, где критически важно минимизировать количество ложных срабатываний.

В перспективах исследования планируется улучшение алгоритма за счет внедрения реальных данных. Еще одним из немаловажных направлений дальнейшей работы является исследование дополнительных каналов обнаружения (акустический, оптический, радиочастотный) и их дальнейшая интеграция в общую комплексную систему защиты от БПЛА.

Литература

1. Дмитриев В.И., Звонарев В.В., Лисицын Ю.Е. Методика обоснования рациональных способов управления беспилотным летательным аппаратом // Труды МАИ. 2020. № 112. URL: http://trudymai.ru/published.php?ID=116566 DOI**:** 10.34759/trd-2020-112-16
2. Seongha Park, Sangmi Shin, Yongho Kim and Eric T. Matson. Combination of Radar and Audio Sensors for Identification of Rotor-type Unmanned Aerial Vehicles (UAVs) DOI: 10.1109/ICSENS.2015.7370533
3. В.М. Карташов, «Методы обнаружения-распознавания радиолокационных, акустических, оптических и инфракрасных сигналов беспилотных летательных аппаратов». Радиотехника, 2021.
4. Балакин, Д.А. Спектральный анализ сигналов в математической среде моделирования Engee: учеб. пособие / Д.А. Балакин, А.А. Кальщиков, Е.В. Шалимова. – М.: Издательство МЭИ, 2024. – 64 с.
5. Семенов В.О., Балакин Д.А. «Основы модельно-ориентированного програмирования в среде Engee на примере моностатического импульсного радара». Нижний Новгород.: Издательство Нижегородского университета им Н. И. Лобачевского.
6. Шестакова, Г.С. «Строение крыльев и механика полета птиц» [Текст]. - Москва: Наука, 1971. - 179 с.
7. Chen, V. C., The Micro-Doppler Effect in Radar, Artech House, 2011
8. M. Tan and Q. Le, "EfficientNetV2: Smaller Models and Faster Training," *arXiv preprint arXiv:2104.00298*, 2021. [Online]. Available: https://arxiv.org/abs/2104.00298.
9. G. Hinton, O. Vinyals, and J. Dean, "Distilling the Knowledge in a Neural Network," *arXiv preprint arXiv:1503.02531*, 2015. [Online]. Available: https://arxiv.org/abs/1503.02531.

IMPROVING THE EFFICIENCY OF UAV DETECTION BASED ON RADAR DATA USING NEURAL NETWORK METHODS

Balakin D.A.[1], Semichastnov A.E.[1], Gorbunov A.R.[1], Akmalov M.R.[1]

[1] ETMC EXPONENTA LTD

The study is dedicated to enhancing the effectiveness of unmanned aerial vehicle (UAV) detection based on radar data using neural network methods. The relevance of the research is driven by the problem of false alarms in detection systems due to the similarity in size between UAVs and birds. To address this challenge, a classification method based on the analysis of micro-Doppler signatures (MDS) is proposed, which reflects the dynamics of object motion. Radar signal modeling was conducted in the Engee environment, where digital models of UAVs and birds were developed. A synthetic dataset comprising several thousand images for each class, including MDS spectrograms, was generated. For classification, EfficientNet architectures (B0 and B7) were utilized, along with a knowledge distillation method. The research demonstrates the potential of integrating neural network algorithms into critical infrastructure protection systems, ensuring the minimization of false alarms. Future prospects of the work include expanding the database with real-world signals and integrating comprehensive detection methods.

ПОВЫШЕНИЕ ЭФФЕКТИВНОСТИ РАСПОЗНАВАНИЯ ОБЪЕКТОВ СИСТЕМОЙ ТЕХНИЧЕСКОГО ЗРЕНИЯ НАЗЕМНОГО РОБОТОТЕХНИЧЕСКОГО КОМПЛЕКСА ЗА СЧЕТ КОМПЛЕКСНОЙ ОБРАБОТКИ ОПТИЧЕСКИХ И РАДИОЛОКАЦИОННЫХ ИЗМЕРЕНИЙ

ст.преп. Безумнов Д.Н., зав.каф. дтн. проф. Чиров Д.С.

Московский технический университет связи и информатики

Аннотация. Одним из важнейших элементов наземного робототехнического комплекса является система технического зрения. С учётом требований к повышению автономности робототехнических комплексов важной функцией робототехнического комплекса является обнаружение и распознавание объектов в поле зрения. Статья посвящена решению задачи повышения эффективности распознавания объектов системой технического зрения за счет комплексирования разнородной информации (оптических и радиолокационных изображений). В качестве распознающего устройства использовалась сверточная нейронная сеть. Показано, что комплексирование информации позволяет повысить эффективность СТЗ РТК на 38-43%.

Введение

Система технического зрения (СТЗ) является важнейшим элементом современных робототехнических комплексов (РТК). В подавляющем большинстве случаев СТЗ передаёт телевизионные (тепловизионные) изображения среды функционирования, в лучшем случае, предварительно обработанные для улучшения качества изображения или стереоизображения.

Однако во многих случаях телевизионной и даже стереотелевизионной информации оказывается недостаточно для эффективного анализа и оценки окружающей обстановки. Наиболее перспективными направлениями исследований в области обнаружения и распознавания объектов, является изучение возможности объединения (комплексирования) информации, полученной из источников на основе различных физических принципов, при этом целью такой операции является получение наиболее полной необходимой информации об объекте исследования [1-5].

Анализ исследований в предметной области

В частности, в работе [1] предлагается комбинированный подход для точной оценки местоположения, размера, позы и информации о движении объекта на основе слияния наблюдений, полученных от стереокамеры и радара, работающего на частоте 24 ГГЦ и регистрирующего до 15 пар наблюдений по дальности и азимуту с идентификационной информацией при каждом сканировании. Предложенный алгоритм прошёл тестирование на синтетических данных и в ходе натурного эксперимента, где движение осуществлялось в дневное и ночное время, в условиях ясной видимости, дождя и снега. Анализ данных после тестирования показал, что все 8 обнаруженных объектов, основанных на алгоритме слияния, на самом деле находились на критическом расстоянии <= 250 мс до столкновения с тестируемым автомобилем, в то время как другое обнаружение было вызвано шумом при прогнозировании траектории объектов, которые, как показал анализ, находились дальше от тестируемого автомобиля.

В работе [5] систематизированы известные технологии комплексирования изображений различных диапазонов и выделены соответствующие им модели, методы и средства. Большое распространение получили системы, снабженные тепловизионной и телевизионной камерами, однако, встречаются и другие. Авторы [5] выделили пять видов комплексирования в зависимости от применяемых методов: работа с яркостями пикселей [6]; замещение фрагментов [7]; критериальное суммирование пикселей [8]; чередование пикселей в результирующем изображении [9]; чередование со смешиванием пикселей в результирующем изображении [10].

В работе [11] предлагается использовать комплексирование данных радара и камеры для обнаружения беспилотным автомобилем соседних автомобилей в потоке. Уникальность работы авторы обуславливают тем, что они не пытаются использовать один сенсор для того, чтобы подтвердить данные второго сенсора, как делали предыдущие авторы, - они определяют трёхмерную позицию объектов, комбинируя мощности двух сенсоров. В системе радар обнаруживает объекты, алгоритм находит эти объекты на кадре видеокамеры, по координате у

определяет их дальность (т.к. на дороге чем автомобиль дальше, тем выше он в кадре), по размеру изображения определяет габариты автомобиля.

Таким образом, анализ проведённых работ показывает, что комплексирование информации для решения задач распознавания объектов СТЗ РТК является перспективным направлением научных исследований в данной области.

Результаты моделирования

В целях оценки эффективности данного подхода рассмотрена гипотеза: комплексирование информации, получаемой РТК через оптический и радиолокационный каналы СТЗ, позволит повысить эффективность, т.е. долю корректно распознанных объектов (КРО).

Информация об объекте, получаемая РТК через оптический канал, представляет собой RGB-изображение размером 32x32 пиксела.

Информация об объекте, получаемая РТК через радиолокационный канал, представляет собой двоичное изображение размером 16x16 пикселов.

Распознавание объектов осуществляется посредством предварительно обученной свёрточной нейронной сети (СНС), модель которой представлена на рисунке 1.

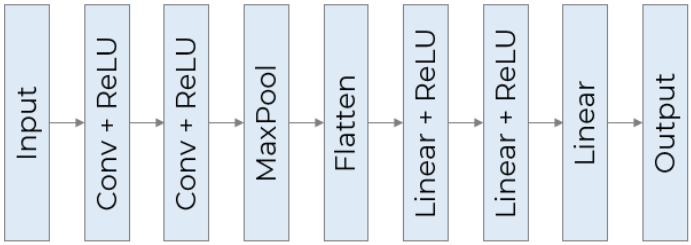


Рисунок 1 – Архитектура СНС для распознавания объектов

СНС реализована средствами фреймворка PyTorch, предобработка данных и проведение экспериментов выполнялось при помощи фреймворков Numpy, Pandas и Scikit-Learn.

Оптические изображения хранятся в памяти БВУ в виде трёхмерных массивов A_i с размерностью 32x32x3, значения которых являются целыми числами в диапазоне от 0 до 255.

Радиолокационные изображения хранятся в памяти БВУ в виде двумерных массивов Bi с размерностью 16x16, элементы которых могут принимать значения 0 или 1.

В целях аугментации данных, т.е. искусственного увеличения объёма обучающего набора, к оптическим и радиолокационным изображениям на входе СНС применяется с вероятностью 0,5 зеркальное отображение по горизонтали.

При подаче на вход СНС массивы A_i и B_i преобразуются в тензоры с размерностью 3x32x32 и 1x16x16 соответственно. Значения массивов A_i также проходят нормировку к диапазону значений [0, 1].

Проведены эксперименты по обучению СНС распознаванию и классификации объектов (препятствий, возникающих на пути наземного мобильного РТК) на оптических и радиолокационных изображениях (см. примеры на рисунке 2) с параметрами, описанными в таблице 1.



а) б)

Рисунок 2 - Примеры оптических (а) и радиолокационных изображений (б)

Таблица 1. Параметры экспериментов по обучению СНС

№ п/п	Параметр	Эксперимент № 1	Эксперимент № 2
1.	Канал получения изображения	Оптический	Радиолокационный
2.	Количество классов	5	5
3.	Объём обучающего набора (количество изображений каждого класса)	450 (90 x 5)	450 (90 x 5)
4.	Объём тестового набора (количество изображений каждого класса)	50 (10 x 5)	50 (10 x 5)
5.	Вероятность зеркального отображения по горизонтали	0,5	0,5
6.	Число эпох обучения	[1, 2, …, 160]	[1, 2, …, 160]
7.	Фиксируемый параметр	Доля КРО, %	Доля КРО, %

В целях повышения информативности графиков, значения доли КРО сглажены методом двустороннего скользящего среднего с шириной окна k = 5:

$$MA_t = \frac{1}{2k+1}\sum_{i=-k}^{k} x_{t+i} = \frac{1}{11}\sum_{i=-5}^{5} x_{t+i}. \quad (1)$$

Комплексирование вероятности КРО по двум каналам выполнено по формуле определения вероятности двух независимых событий:

$$P_К = P_{ОЭ} + P_{РЛ} - P_{ОЭ} \cdot P_{РЛ}, \quad (2)$$

где: $P_{ОЭ}$ – вероятность (доля КРО) для оптического канала, $P_{РЛ}$ – вероятность (доля КРО) для радиолокационного канала.

Результаты экспериментов № 1 и № 2 и вероятность КРО, вычисленная по формуле (2), представлены на рисунке 3.

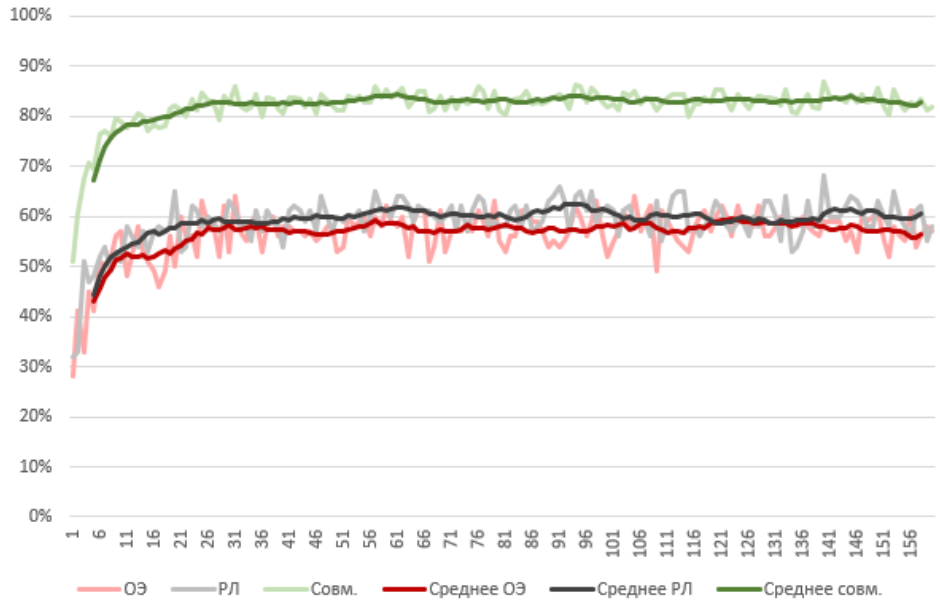


Рисунок 3 – Зависимость доли КРО от числа эпох обучения

Как видно из представленного графика, существенный рост доли КРО прекращается при 60 эпохах обучения. Средняя доля КРО для числа эпох обучения в диапазоне от 60 до 160 составляет: 58% для оптического канала, 60% для радиолокационного канала, 83% для комплексированной оценки КРО.

Заключение

Таким образом, принятие решения о принадлежности объекта к определённому классу на основании информации, полученной через оптический и радиолокационный каналы, позволяет

повысить эффективность СТЗ РТК на 43% по сравнению с использованием только оптического канала и на 38% по сравнению с использованием только радиолокационного канала.

Литература

1. Shunguang Wu, S. Decker, P. Chang, T. Camus and J. Eledath, "Collision sensing by stereo vision and radar sensor fusion," 2008 IEEE Intelligent Vehicles Symposium, Eindhoven, Netherlands, 2008, pp. 404-409, doi: 10.1109/IVS.2008.4621137.
2. Чиров Д.С., Чертова О.Г., Потапчук Т.Н. Методика обоснования требований к системе технического зрения робототехнического комплекса // Труды СПИИРАН. – № 2(51). – 2017. – С. 152-176. DOI: 10.15622/sp.51.7. EDN: YIASIV.
3. Чиров Д.С., Терешонок М.В., Елсуков Б.А. Метод и алгоритмы оптимизации технических характеристик комплексов радиомониторинга // TComm. – №10. – 2014. – С. 88-92. EDN: SZZOWP.
4. Чиров Д.С., Хрипунов С.П. Интеллектуальные методы распознавания тактических ситуаций в условиях автономного применения робототехнических комплексов военного назначения // Вопросы безопасности. – 2017. – № 1. – С. 22-34. DOI: 10.7256/2409-7543.2017.1.21643. EDN: XXRDAX.
5. Рогачёв В.А., Ципуринда С.Л. Исследование методов комплексирования изображений различных диапазонов // Труды учебных заведений связи. – 2016. – № 3. – С. 59-64.
6. Бендицкий А. А. Способ формирования матричного объекта. Патент № 2305320 РФ; заявл. 02.07.2004 № 2004120046/09; опубл. 27.08.2007. Бюл. 24.
7. Кузнецов А.Е., Калюжный В.И., Ковалев А.О., Ефремов И.Ф., Гектин Ю.М. Способ дистанционного обследования объектов электрических сетей с помощью тепловидеосъемочного устройства. Патент № 2258204 РФ; заявл. 16.02.2004 № 2004104193/28; опубл. 10.08.2005. Бюл. 22.
8. Богданов А.П., Костяшкин Л.Н., Морозов А.В., Павлов О.В., Романов Ю.Н., Рязанов А.В. Способ комплексирования цифровых полутоновых телевизионных и тепловизионных изображений. Патент № 2451338 РФ; заявл. 23.12.2010 № 2010152858/08; опубл. 20.05.2012 Бюл. № 14.
9. Ерош И.Л., Соловьев Н.В., Филатов Г.П., Сергеев М.Б., Козлов А.А., Литвинов М.Ю. Способ совмещения изображений, полученных с помощью различных фотодатчиков и устройство для его реализации. Патент № 2435221 РФ; заявл 10.12.2007. № 2007145889/28; опубл. 27.11.2011 Бюл. № 33.
10. Ветров А.Н., Осипова А.А., Гахзар М.А., Артюхин И.Ю. Совмещение изображений от датчиков различных диапазонов частот // Современные проблемы науки и образования: электрон. научн. журн. 2013. № 6. URL: http://www.scienceeducation.ru/ru/article/view?id=11589 (дата обращения 27.12.2024).
11. A. Sikdar, S. Cao, Y.F. Zheng and R.L. Ewing, "Radar depth association with vision detected vehicles on a highway," 2014 IEEE Radar Conference, Cincinnati, OH, USA, 2014, pp. 1159-1164, doi: 10.1109/RADAR.2014.6875771.

INCREASING THE EFFICIENCY OF OBJECT RECOGNITION BY THE VISION SYSTEM OF A GROUND-BASED ROBOTIC COMPLEX BY INTEGRATED PROCESSING OF OPTICAL AND RADAR MEASUREMENTS

Senior Lecturer Bezumnov D.N., Head of Department, Dr. Sci. (Eng.), Prof. Chirov D.S.

Moscow Technical University of Communications and Informatics

Abstract. One of the most important elements of a ground-based robotic complex is a vision system. Taking into account the requirements for increasing the autonomy of robotic complexes, an important function of a robotic complex is the detection and recognition of objects in the field of view. The article is devoted to solving the problem of increasing the efficiency of object recognition by a vision system by integrating heterogeneous information (optical and radar images). A convolutional neural network was used as a recognition device. It has been shown that information integration allows increasing the efficiency of the RTK STS by 38-43%.

АНСАМБЛИРОВАНИЕ МЕТРИК КАЧЕСТВА ИЗОБРАЖЕНИЙ ЛИЦ ДЛЯ ПОВЫШЕНИЯ СООТВЕТСТВИЯ ЭКСПЕРТНЫМ ОЦЕНКАМ

асп. Бордюжа В., проф. Умняшкин С.В.

Национальный исследовательский университет «МИЭТ»

Оценка качества изображений лиц — это комплексная и сложная задача, требующая учёта множества факторов. Успешное решение этой задачи играет ключевую роль в таких областях, как распознавание лиц, генерация и восстановление изображений. В данной работе предлагается подход, основанный на ансамблировании и тонкой настройке предобученных нейросетевых моделей, который позволяет компенсировать недостатки отдельных методов. В качестве базовых алгоритмов были выбраны PaQ-2-PiQ, HyperIQA и WaDIQaM. Ансамбль строился с учётом дообучения на наборах данных CGFIQA-40k и GFIQA-20k. Результаты экспериментов показывают, что предложенный метод ансамблирования обеспечивает более точную оценку качества изображений лиц по сравнению с такими классическими подходами к построению ансамблей, как бустинг и стекинг. В работе проведено сравнение предложенного метода с современными нейросетевыми подходами для оценки качества изображений лиц, такими как StyleGAN-IQA и DSL-FIQA. Тестирование выполнялось на наборе данных PIQ23, который не участвовал в обучении моделей, что позволило объективно оценить их способность к обобщению на новых данных. Согласно результатам экспериментов, нейросетевой ансамбль превосходит StyleGAN-IQA и DSL-FIQA по показателю коэффициента корреляции Пирсона, демонстрируя более высокую согласованность с экспертными оценками.

Введение

Оценка качества изображений человеческих лиц играет важную роль в различных задачах компьютерного зрения, таких как распознавание лиц [1], генерация изображений и их восстановление [2], а также в системах видеонаблюдения, биометрической идентификации, контроле доступа и аналитике видеопотоков в реальном времени. В этих областях качество изображения напрямую влияет на надежность и точность последующих этапов обработки: от детекции и трекинга до финальной идентификации личности [3]. Для достижения высокой точности алгоритмов требуется объективная и согласованная с человеческим восприятием метрика качества изображений, которая способна учитывать особенности данных, описывающих лица. Разработка эффективной метрики для оценки качества изображения лица сопряжена с преодолением значительных трудностей. Присущая человеческим лицам сложность, выражающаяся в тонких визуальных особенностях и разнообразии мимических выражений, оказывает значительное влияние на воспринимаемое качество изображений. Эти проблемы осложняются окклюзиями лица, вызванными масками и аксессуарами, а также ограниченностью доступных лицензионных датасетов.

Существующие подходы к оценке качества изображений лиц можно условно разделить на две категории: общие методы оценки качества изображений и методы оценки качества для биометрии. Общие методы оценки качества изображений, такие как DBCNN [4], PaQ-2-PiQ [5], HyperIQA [6], демонстрируют стабильные результаты на наборах данных KonIQ-10k [7] и LIVEC [8], но часто игнорируют специфику, характерную для лиц. В подходах, направленных на оценку качества для биометрических целей, идентификация достигается за счёт использования таких важных для лица факторов, как антропометрия и освещение. Однако такие методы не оценивают, насколько изображение выглядит "качественным" с точки зрения человеческого восприятия (например, насколько изображение кажется естественным или приятным глазу).

Целью данной работы является повышение корреляции между метриками качества изображений лиц и человеческими оценками путем использования методов ансамблирования метрик и дальнейшего развития подходов, представленных в статье [9].

Предлагаемые способы ансамблирования метрик качества изображений

Как уже было показано в работе [9], использование ансамбля из метрик качеств позволяет в значительной степени улучшить корреляцию между предсказанными значениями и экспертными оценками качества. Предложенные в работе [9] методы ансамблирования обладают рядом недостатков. Ввиду того, что использовались уже обученные нейросетевые модели, возникает некоторая избыточность в количестве используемых алгоритмов. Для компенсации недостатков одного алгоритма приходится использовать несколько других, что в итоге, безусловно, позволяет улучшить качество предсказаний. Можно предположить, что для

ансамблирования можно использовать меньшее количество алгоритмов, если иначе использовать идею построения ансамбля.

В качестве базовых нейросетевых моделей для безэталонной оценки качества изображений мы выбрали алгоритмы PaQ-2-PiQ [5], HyperIQA [6] и WaDIQaM [10].

Авторы алгоритма PaQ-2-PiQ использовали предобученную на наборе данных ImageNet [12] модель ResNet-18 [11] из фреймворка PyTorch [13] и сохранили только её сверточную часть для последующего обучения. На основе этой структуры добавили два слоя пулинга (адаптивный усредняющий пулинг и адаптивный максимальный пулинг) и два полносвязных слоя. Последний полносвязный слой выдавал единственное число – оценку качества изображения.

Архитектура глубокой нейронной сети WaDIQaM состоит из 10 сверточных слоев, 5 слоёв субдискретизации, а также 2 полносвязных слоёв, что делает ее значительно более глубокой по сравнению с другими моделями оценки качества, доступными на момент подготовки авторами WaDIQaM их работы. Уникальность предложенной архитектуры состоит в том, что с небольшими адаптациями она может быть использована как для эталонной, так и для безэталонной оценки качества изображений.

Авторы HyperIQA предложили самоадаптивную гибридную сеть для безэталонной оценки качества изображений в естественных условиях. Архитектура включает три этапа:

1. Понимание содержания – извлечение многомасштабных признаков с помощью модифицированной архитектуры ResNet50 [11].
2. Обучение правилам восприятия – гиперсеть генерирует веса и смещения для целевой сети на основе извлеченных признаков.
3. Прогнозирование качества – целевая сеть из полносвязных слоев преобразует признаки в оценку качества.

Такой подход позволяет учитывать как содержание изображения, так и правила восприятия для более точного предсказания качества.

Для решения задачи ансамблирования мы предлагаем воспользоваться тонкой настройкой (дообучением) нейросетевых алгоритмов. Вместо того, чтобы обучать их с нуля со случайной инициализацией весов, мы заморозим большую часть весов для каждой из архитектур, взяв предобученные веса для набора данных KonIQ-10k [7]. Для агрегации предсказаний выходные значения из нейросетевых моделей подаются на вход в нейрон с линейной функцией активации (рисунок 1). Несмотря на то, что оценка качества восприятия лица требует более специфичных признаков, чем оценка качества изображения в целом, на первых слоях нейронной сети извлекаются низкоуровневые признаки, которые должны быть схожи вне зависимости от типа задачи оценки качества. Возможность дообучаться должна позволить алгоритмам не только компенсировать недостатки друг друга, но также превзойти показатели, доступные на этапе первоначального обучения, за счет адаптации к новым данным. У нейросетевого алгоритма WaDIQaM будет происходить дообучение двух полносвязных слоев. В случае с алгоритмом HyperIQA будет заморожена только сеть для извлечения признаков (ResNet50), тогда как остальные компоненты архитектуры будут дообучены. Аналогичный подход будет применен и к архитектуре PaQ-2-PiQ.

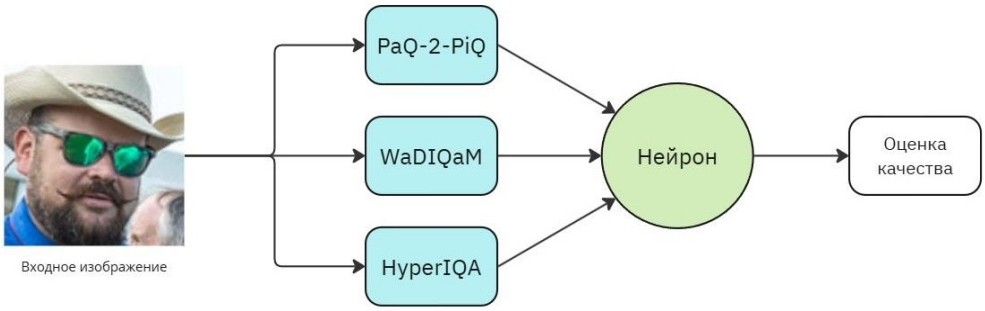


Рисунок 1 – Архитектура предложенного метода ансамблирования

Используемые наборы данных для оценки качества

Для оценки качества ансамбля мы использовали три набора данных: GFIQA-20k [14], PIQ23 [15] и CGFIQA-40k [16].

До недавнего времени модели общей оценки качества лица тестировались только на наборах данных PIQ23 и GFIQA-20k. Набор PIQ23 включает в себя разнообразные изображения из реальной среды с некадрированными лицами, но его небольшой объём ограничивает эффективность обучения устойчивых моделей. Кроме того, оба набора данных имеют дисбаланс в распределении полов и оттенков кожи (таблица 1). Это несоответствие может привести к появлению предвзятости при обучении моделей, что снижает их производительность и надежность при оценке качества изображений с лицами. Предыдущие исследования [17, 18] показали, что такой дисбаланс в распределении данных оказывает значительное негативное влияние на эффективность моделей в различных приложениях, связанных с обработкой лиц.

Таблица 1. Сравнение наборов данных по размеру выборки, цвету кожи и распределению по полу [16]

Набор данных	Количество образцов	Оттенок кожи, %			Пол, %	
		Светлый	Средний	Темный	Мужской	Женский
PIQ23	5116	74.5	9.0	16.5	94.0	6.0
GFIQA-20K	20000	81.6	6.7	11.7	64.2	35.8
CGFIQA-40k	39312	53.83	24.91	21.26	51.50	48.50

Для решения этой проблемы был создан набор данных под названием CGFIQA-40k [16], который включает в себя около 40000 изображений, каждое с разрешением 512x512 пикселей. Для каждого изображения было собрано по 20 оценок от привлечённых участников, однако информация об их квалификации и принципах отбора в публикации отсутствует [16]. Из первоначального набора в 40000 изображений его создатели отфильтровали небольшое количество изображений с непригодным содержанием или неполными метками, в результате чего получилось 39312 качественных изображений. Этот набор данных специально подобран таким образом, чтобы включить в себя обширную коллекцию изображений лиц с различным распределением по оттенку кожи, полу и наличию на лице окклюзий, таких как маска и аксессуары (рисунок 2).

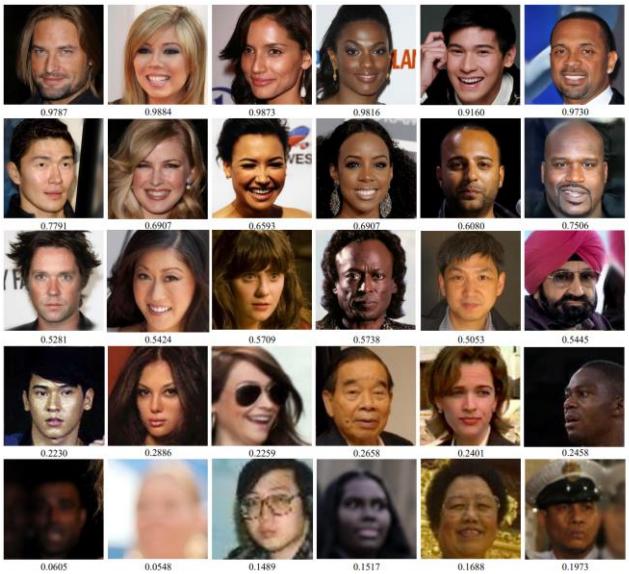


Рисунок 2 – Примеры лиц из набора данных CGFIQA-40k, отражающие разнообразие визуального качества по пяти категориям: отличное (0.8,1], хорошее (0.6,0.8], приемлемое (0.4,0.6], низкое (0.2,0.4] и неудовлетворительное [0,0.2]

Результаты экспериментов и выводы

Результаты работы оцениваются с помощью коэффициента линейной корреляции Пирсона (PLCC) и коэффициента ранговой корреляции Спирмена (SRCC) между экспертными оценками качества изображения и результатом автоматизированного предсказания алгоритма.

Для обучения, тестирования и валидации мы использовали официальные разбиения наборов данных GFIQA-20k и CGFIQA-40k от их авторов.

Предложенный метод ансамблирования сравнивался с двумя современными нейросетевыми подходами StyleGAN-IQA [14] и DSL-FIQA [16]. Кроме того, для честного сравнения мы дообучали каждую из нейронных сетей ансамбля по отдельности, используя методику, аналогичную описанной ранее. Помимо этого, мы протестировали подходы к ансамблированию, предложенные в работе [9], где использовались линейная регрессия и бустинг на деревьях. В качестве базовых алгоритмов выступали дообученные по отдельности нейронные сети. Для оценки обобщающей способности моделей мы проверили их на наборе данных PIQ23, который не использовался ни на этапе обучения, ни на этапе валидации, что позволило объективно оценить способность работать с новыми данными.

Результаты тестирования для коэффициента корреляции Пирсона и ранговой корреляции Спирмена приведены в таблице 2.

Таблица 2. Значение коэффициентов корреляции между экспертной и машинной оценкой для наборов данных GFIQA-20k и CGFIQA-20k

Тестируемый алгоритм	**GFIQA-20k**		**CGFIQA-40k**	
	PLCC	SRCC	PLCC	SRCC
StyleGAN-IQA	*0.967*	*0.968*	*0.982*	*0.982*
DSL-FIQA	**0.975**	**0.974**	**0.987**	**0.988**
PaQ-2-PiQ (без дообучения)	0.759	0.770	0.853	0.890
HyperIQA (без дообучения)	0.896	0.903	0.892	0.939
WaDIQaM (без дообучения)	0.679	0.685	0.675	0.551
PaQ-2-PiQ	0.885	0.890	0.950	0.947
HyperIQA	0.907	0.927	0.926	0.959
WaDIQaM	0.733	0.742	0.894	0.886
Нейросетевой ансамбль	0.954	0.955	0.981	0.980
Бустинг на деревьях [19]	0.931	0.933	0.968	0.965
Линейная регрессия [20]	0.926	0.932	0.958	0.959

Как следует из приведённых в таблице результатов, нейросетевой метод ансамблирования лучше справляется с задачей оценки качества лица, демонстрируя более высокую точность и стабильность результатов по сравнению с альтернативными подходами к построению ансамблей. Важно отметить, что хотя значения корреляции нейросетевых моделей StyleGAN-IQA и DSL-FIQA выше, чем у предложенного нейросетевого ансамбля, их количество параметров превышает количество параметров нашего ансамбля в 1.8 и 5.7 раз соответственно (таблица 3). Это приводит к повышенным требованиям к вычислительным ресурсам как на этапе обучения, так и при инференсе, что, соответственно, может затруднять их использование в вычислительно ограниченных системах и при решении задач, требующих обработки данных в реальном времени.

Таблица 3. Сравнение количества параметров и обучаемых параметров моделей

Алгоритм	Количество параметров	Обучаемые параметры
PaQ-2-PiQ	1.2×10^7	0.5×10^6
HyperIQA	2.7×10^7	3.9×10^6
WaDIQaM	5.2×10^6	0.5×10^6
Нейросетевой ансамбль	4.4×10^7	4.9×10^6
StyleGAN-IQA	8.0×10^7	5.4×10^7
DSL-FIQA	2.5×10^8	2.5×10^8

Несмотря на значительную разницу в количестве обучаемых параметров (таблица 3), нейросетевой ансамбль превосходит алгоритм StyleGAN-IQA по показателям корреляции, независимо от используемого набора данных для обучения. Использование ансамбля улучшило

значение коэффициента корреляции Пирсона (таблица 4) между экспертными оценками и предсказанными значениями модели относительно других рассматриваемых алгоритмов.

Таблица 4. Значение коэффициента корреляции Пирсона и Спирмена экспертных оценок с результатом автоматизированного метода оценивания качества изображений лиц для PIQ23

Алгоритм	Набор данных для обучения	PLCC	SRCC
StyleGAN-IQA	**GFIQA-20k**	0.332	0.342
StyleGAN-IQA	**CGFIQA-40k**	0.354	0.364
DSL-FIQA	**GFIQA-20k**	0.395	*0.417*
DSL-FIQA	**CGFIQA-40k**	*0.423*	**0.465**
Нейросетевой ансамбль	**GFIQA-20k**	0.406	0.395
Нейросетевой ансамбль	**CGFIQA-40k**	**0.446**	0.398

Согласно проведенным экспериментам, нейросетевое ансамблирование является более предпочтительным подходом в сравнении с классическими методами построения ансамблей. Предполагается, что увеличение числа разнообразных алгоритмов в ансамбле должно положительно повлиять на значение коэффициента корреляции между автоматизированными и экспертными оценками качества изображений человеческих лиц.

Литература

1. CR-FIQA: Face image quality assessment by learning sample relative classifiability / Boutros A., Fang M., Klemt M., Fu B., Damer N. // IEEE Conference on Computer Vision and Pattern Recognition. 2023. P. 5836–5845.
2. VQFR: Blind face restoration with vector-quantized dictionary and parallel decoder / Gu Y., Wang X., Xie L., Dong C., Li G., Shan Y., Cheng M.-M. // European Conference on Computer Vision. 2022. P. 126–143.
3. SDD-FIQA: Unsupervised face image quality assessment with similarity distribution distance / Ou F.-Z., Chen X., Zhang R., Huang Y., Li S., Li J., Li Y., Cao L., Wang Y.-G. // IEEE Conference on Computer Vision and Pattern Recognition. 2021. P. 7670–7679.
4. Blind image quality assessment using a deep bilinear convolutional neural network / Zhang W., Ma K., Yan J., Deng D., Wang Z. // IEEE Transactions on Circuits and Systems for Video Technology. 2020. Vol. 30. No. 1. P. 36–47.
5. From patches to pictures (PaQ-2-PiQ): Mapping the perceptual space of picture quality / Ying Z., Niu H., Gupta P., Mahajan D., Ghadiyaram D., Bovik A. // IEEE Conference on Computer Vision and Pattern Recognition. 2020. P. 3575–3585.
6. Blindly assess image quality in the wild guided by a self-adaptive hyper network / Su S., Yan Q., Zhu Y., Zhang C., Ge X., Sun J., Zhang Y. // IEEE Conference on Computer Vision and Pattern Recognition. 2020. P. 3667– 3676.
7. Lin H., Hosu V., Saupe D. Koniq-10K: Towards an ecologically valid and large-scale IQA database // arXiv.org. 2018. URL: https://arxiv.org/pdf/1803.08489.pdf (дата обращения: 11.01.2024).
8. Ghadiyaram D., Bovik A. C. Massive online crowdsourced study of subjective and objective picture quality // IEEE Transactions on Image Processing. 2016. Vol. 25. No. 1. P. 372–387.
9. Бордюжа, В. Ансамблирование метрик качества изображений для повышения корреляции с экспертными оценками / В. Бордюжа, С. В. Умняшкин // Цифровая обработка сигналов и ее применение DSPA - 2024 : Доклады XXVI Международной конференции, Москва, 27–29 марта 2024 года. – Москва: Российское научно-техническое общество радиотехники, электроники и связи им. А.С. Попова, 2024. – С. 187-192.
10. Deep Neural Networks for No-Reference and Full-Reference Image Quality Assessment / Bosse S., Maniry D., Müller K.-R., Wiegand T., Samek W. // IEEE Transactions on Image Processing. 2017. Vol. 27. P. 206-219.
11. Deep Residual Learning for Image Recognition / Kaiming He, Zhang X., Ren S., Sun J. // IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2016. P. 770–778.

12. ImageNet large scale visual recognition challenge / Russakovsky O., Deng J., Su H., Krause J., Satheesh S., Ma S., Huang Z., Karpathy A., Khosla A., Bernstein M., Berg A. C., Fei-Fei L. // International Journal of Computer Vision (IJCV). 2015. Vol. 115. No. 3. P. 211–252.

13. PyTorch: An open-source machine learning framework // PyTorch [Электронный ресурс]. URL: https://pytorch.org (дата обращения: 23.01.2025).

14. Going the Extra Mile in Face Image Quality Assessment: A Novel Database and Model / Su S., Lin H., Hosu V., Wiedemann O., Sun J., Zhu Y., Liu H., Zhang Y., Saupe D. // IEEE Transactions on Multimedia. 2023. P. 2671–2685.

15. An Image Quality Assessment Dataset for Portraits / Chahine N., Calarasanu S., Garcia-Civiero D., Cayla T., Ferradans S., Ponce J. // IEEE Conference on Computer Vision and Pattern Recognition. 2023. P. 9968–9978.

16. DSL-FIQA: Assessing Facial Image Quality via Dual-Set Degradation Learning and Landmark-Guided Transformer / Chen W.-T., Krishnan G., Gao Q., Kuo S.-Y., Ma S., Wang J. // IEEE Conference on Computer Vision and Pattern Recognition. 2024. P. 2931–2941.

17. Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification / Buolamwini J., Gebru T. // Proceedings of the 1st Conference on Fairness, Accountability and Transparency. PMLR. 2018. Vol. 81. P. 77–91.

18. Demographic Effects in Facial Recognition and Their Dependence on Image Acquisition: An Evaluation of Eleven Commercial Systems / Cook C. M., Howard J. J., Sirotin Y. B., Tipton J. L., Vemury A. R. // IEEE Transactions on Biometrics, Behavior, and Identity Science. 2019. Vol. 1. No.1. P. 32–41

19. Catboost: unbiased boosting with categorical features / Prokhorenkova L., Gusev G., Vorobev A., Dorogush A. V., Gulin A. // 32nd Conference on Neural Information Processing Systems (NeurIPS). 2018. P. 6638-6648.

20. Scikit-learn: Machine Learning in Python // Scikit-learn [Электронный ресурс]. URL: https://scikit-learn.org/stable (дата обращения: 10.01.2024).

ENSEMBLING FACE IMAGE QUALITY METRICS TO IMPROVE CONSISTENCY WITH EXPERT EVALUATIONS

Bordiuzha V., Umnyashkin S.V.

National Research University of Electronic Technology

Face image quality assessment is a complex and challenging task that requires consideration of many different factors. It plays a major role in such areas as face recognition, image generation and reconstruction. In this paper we propose an approach based on the ensembling and fine-tuning of pre-trained neural network models, which compensates disadvantages of individual methods for face image quality assessment. PaQ-2-PiQ, HyperIQA and WaDIQaM were chosen as the basic algorithms. The ensemble built was pre-trained on the CGFIQA-40k and GFIQA-20k datasets. Experimental results show that the proposed ensembling method provides more accurate estimation of face image quality as compared with classical ensemble building approaches such as boosting and stacking. The paper compares the proposed method with state-of-the-art neural network approaches for face image quality assessment, such as StyleGAN-IQA and DSL-FIQA. Testing was performed on PIQ23 dataset, which was not included in the training of the models, ensuring fair evaluation on new data. According to the experimental results, the proposed neural network ensemble outperforms both StyleGAN-IQA and DSL-FIQA algorithms in terms of Pearson's correlation, demonstrating better correspondence with expert evaluations.

ПРИМЕНЕНИЕ ПРОГРАММНО-КОНФИГУРИРУЕМОГО РАДИО В КЛАСТЕРНОЙ БЕСПРОВОДНОЙ СЕНСОРНОЙ СЕТИ

асп. Гурьянов Е.Д.

Ярославский государственный университет им. П.Г. Демидова, г. Ярославль, Россия

Аннотация. Представлен анализ перспектив использования программно-конфигурируемого радио (ПКР, SDR – англ. Soft Defined Radio) в качестве приемо-передающей аппаратуры для имитации узла кластерной беспроводной сенсорной сети (БСС) для автономных мобильных платформ. В исследовании рассматривается алгоритм декодирования принятых сообщений в приложении к многоскачковой сенсорной сети с системой связи на базе чипсетов NRF24L01+ (протокол ESB) и радиосети на базе модулей LoRa. Рассматривается возможность оснащения автономных мобильных платформ ПКР приемниками, а также дополнительные возможности, которые открывает применение ПКР в мобильной робототехнике.

Введение

В настоящее время использование программно-конфигурируемого радио [1, 2] применительно к мобильной робототехнике открывает значительные перспективы перед разработчиками. Возможность унификации связного оборудования мобильного робота с уже используемыми системами и сетями радиосвязи, проведение различных оценок состояния радиоэфира с помощью программного обеспечения (ПО), располагающегося на борту робота, приложения для перехвата управления беспилотными летательными аппаратами (БПЛА, БЛА) нарушителей и т.д. В представленном исследовании рассматривается возможность использования ПКР приемника с системой на кристалле (СнК, SoC – англ. System on a Chip) для имитации узла беспроводной сенсорной сети. В перспективе это позволит применить полученные результаты в области как построения аналога узла на базе ПКР, так и узла-агрегатора при объединении нескольких подсетей-кластеров с различными системами радиосвязи.

Обзор технологий СнК и ПКР

Активное развитие технологии ПКР, а также необходимость наличия вычислителя непосредственно на борту устройства (например, БЛА) привело к объединению плат, сочетающих в себе однокристальную систему – систему на кристалле (СнК/SoC), ПЛИС (FPGA) и чипсет АЦП/ЦАП. Однокристальная система – электронная схема, выполняющая функции целого устройства (например, компьютера) и размещенная на одной интегральной схеме. В зависимости от назначения она может оперировать цифровыми или аналоговыми сигналами, а также частотами используемого радиодиапазона. Как правило, подобные устройства применяются в портативных и встраиваемых системах. Одним из примеров таких устройств может являться ПКР с SoC Pluto Plus или LibreSDR [3], сочетающие в себе SoC Zynq 7010 или Zynq 7020 и чипсет АЦП/ЦАП AD9363. Наличие на борту ПКР SoC позволяет развернуть полноценную Linux систему с необходимыми программными пакетами для проведения различных вычислений и работы с сигналами. Кроме того, широкий диапазон АЦП/ЦАП от 70 МГц до 6 ГГц с полосой 20 МГц на канал дает возможность работать с большинством типов сигналов, применяемых в БСС в решениях IoT и связи с беспилотными аппаратами.

В настоящее время в мире ведутся активные работы по развитию беспилотной авиации различного назначения, а также созданию новых БЛА, способных решать задачи в составе группы. Выполнение задач такой группой позволяет повысить эффективность выполнения задачи до 20% [4]. При построении надежной системы связи в гетерогенной группе автономных мобильных платформ, включающей наземный и воздушный сегменты, необходимо учитывать особенности рельефа местности, диапазоны частот радиосвязи экстренных служб и т.п. В определенных условиях возможно также противодействие систем радиоэлектронной борьбы. Для решения таких задач бортовая система радиосвязи должна обладать гибкостью для изменения частоты вещания, модуляции, применения кодирования.

Системы программно-конфигурируемого радио позволяют учесть многие из перечисленных требований.

В приведенном исследовании все эксперименты проводились с помощью ПКР Pluto Plus в режиме 1r1t, что соответствует одному каналу на прием и одному каналу на передачу. Включение более сложного режима 2r2t позволяет задействовать два независимых канала на прием и 2 на передачу, а также работать с технологией MIMO. В программных пакетах, типа GNURadio и Matlab, для работы с ПКР приемниками существует возможность коммуникации с АЦП/ЦАП Pluto Plus напрямую, через USB 2.0. Это упрощает первичный анализ и разработку, однако позволяет работать только в режиме 1r1t, поскольку модуль для коммуникации с устройством написан под ПКР ADALM Pluto, имеющий только 1 канал на прием и 1 на передачу.

Требование внедрения таких устройств в БСС обуславливается необходимостью наличия узлов-агрегаторов в больших сетях, а также узлов-мониторов в критически важных для инфраструктуры сетях или сетях, подверженных атакам извне в силу их важности. В проводимом исследовании ПКР с SoC применяется как узел-агрегатор при соединении кластеров сети с различными типами систем связи. В данном случае используется кластер, построенный на чипсетах NRF24L01+ с модуляцией GMSK, и кластер, состоящий из модулей LoRa с модуляцией ЛЧМ.

Практическая реализация

Рассматривается прием и декодирование сообщений, передающихся в самоорганизующейся радиосети с использованием библиотеки RF24Mesh [5] и библиотеки LoRa [6] с помощью SDR приемника. В конечном итоге все сводится к программной имитации узла сети с помощью ПК, а в дальнейшем – микрокомпьютера типа Raspberry Pi Zero или его аналогов, а также платы SDR Hack RF. Наилучшим вариантом является реализация функционала узла на SDR устройстве Pluto Plus или LibreSDR, что позволяет реализовать узел сети на одной плате. Параметры радиосети следующие: частота 2,481 ГГц, скорость передачи 2 Мбит/с, мощность передатчика 1 мВт. Количество узлов в сети 3: мастер-узел, узел-приёмник, узел-передатчик. Узел-передатчик отправляет сообщения на мастер-узел и узел-приёмник, а мастер-узел осуществляет пингование дочерних узлов. Эксперимент заключается в исследовании возможности приема и декодирования пакетов с помощью SDR приемника и GNU Radio, при этом для декодирования используется утилита из проекта Cyber Explorer [7]. Одновременно в GNU Radio ведется запись оцифрованного эфира. Схема проекта GNU Radio представлена на рисунке 1.

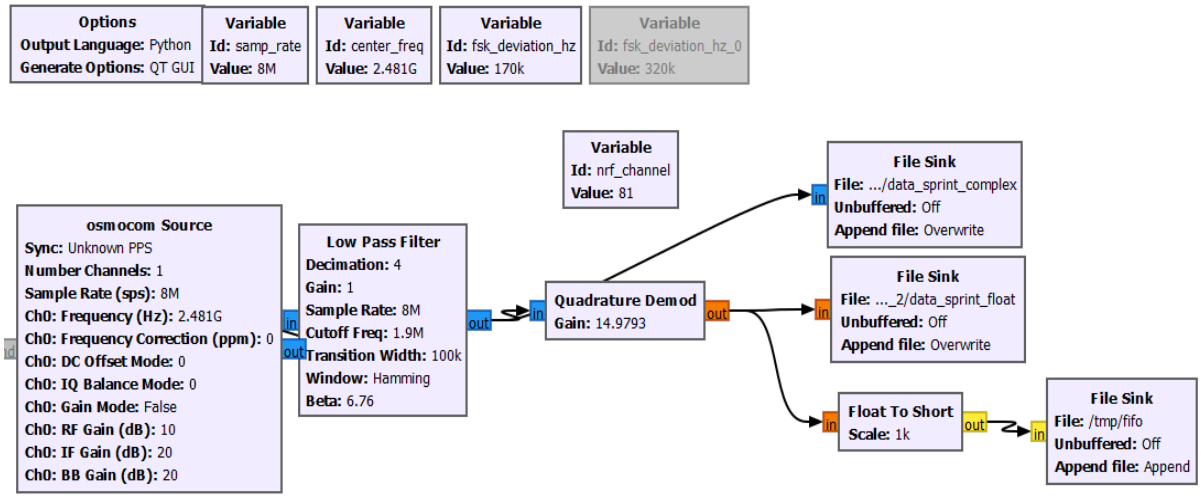


Рисунок 1. Схема приемной части в эксперименте 1 в GNU Radio

Прием и декодирование сигналов узла БСС на базе чипсета NRf24L01+ происходит в он-лайн режиме. Конфигурация сети такова, что узел-передатчик и узел-приемник находятся на одном уровне. Узел-передатчик отправляет на мастер-узел нагрузку в формате uint8 255, 0, 255, 0, 255, 0, 255, 0, 255, 0, что соответствует FF 00 FF 00 FF 00 FF 00 в ОСТ формате. Как можно видеть из рисунка 2, где изображён вывод консоли утилиты декодера, полю data соответствует первый октет – 05 – адрес отправителя, третий октет – 00 – адрес получателя, пятый октет – 93 – порядковый номер сообщения, седьмой октет – 45 – тип сообщения, который в библиотеке RF24Mesh соответствует типу E. Далее, начиная с девятого октета, идет нагрузка. NodeID 5, его адрес 50, т.е. адрес узла фактически 5, 0 – это материнский узел в формате HEX – так реализовано в коде для Arduino, в терминале адрес в ОСТ (слева направо), то есть 05 – тут он совпадает.

```
Address: 0xCCCCCCCCE3 length:18, pid:0, no_ack:0, CRC:0x5130 data:05 00 00 00 8D 00 45 20 FF 00 FF 00 FF 00 FF 00 FF 00
Address: 0xCCCCCCCCE3 length:18, pid:0, no_ack:0, CRC:0x5130 data:05 00 00 00 8D 00 45 20 FF 00 FF 00 FF 00 FF 00 FF 00
Address: 0xCCCCCCCCE3 length:18, pid:0, no_ack:0, CRC:0x5130 data:05 00 00 00 8D 00 45 20 FF 00 FF 00 FF 00 FF 00 FF 00
Address: 0xCCCCCCCCE3 length:18, pid:0, no_ack:0, CRC:0x5130 data:05 00 00 00 8D 00 45 20 FF 00 FF 00 FF 00 FF 00 FF 00
```

Рисунок 2. Декодирование данных с чипсета NRF24L01+

Таким образом, можно сделать вывод о том, что декодер отрабатывает правильно. Конечно, необходима доработка и улучшение программы, для того чтобы привести работу декодера в соответствие с функционалом оригинальной библиотеки RF24Mesh и перенести всю обработку в GNU Radio.

Эксперимент 2. В данном случае прием и декодирование сообщений с радиомодулей LoRa осуществляется с помощью SDR устройства PlutoPlus, а также библиотек gr-lora от rpp0 [8] и gr-lora_sdr от Joachim Tapparel [9] для GNU Radio. Схема сборки исследуемого проекта в GNU Radio представлена на рисунке 3.

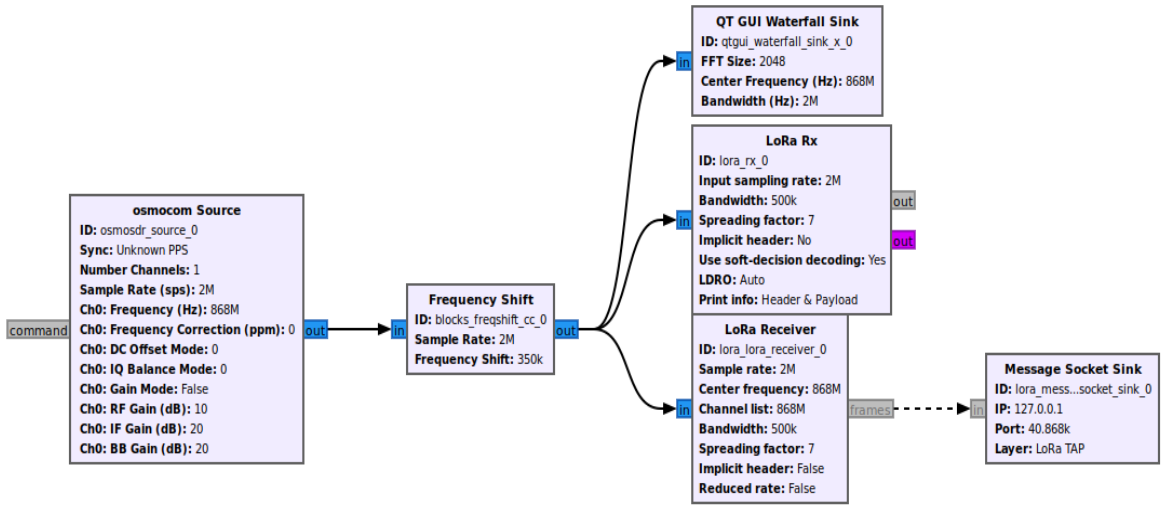


Рисунок 3. Схема каналов обработки в эксперименте 2

Установлено, что обе библиотеки показали одинаковое качество декодирования на «чистом» сигнале (рисунок 4), то есть в эфире сигнал с ЛЧМ модуляцией передается без преднамеренных помех. В некоторых заводских мультикоптерах, применяющих в качестве управляющего канала технологию LoRa, применяется метод расширения спектра ППРЧ для большей скрытности канала. В простом варианте приема сообщений каналы приема могут быть сдвинуты по частотам передачи, в таком случае остается задача лишь объединения сообщений в один поток. Кроме того, следует отметить, что библиотека gr-lora_sdr позволяет также вносить изменения с помощью дополнительных блоков в тракт демодуляции и декодирования, что представлено на рисунке 5.

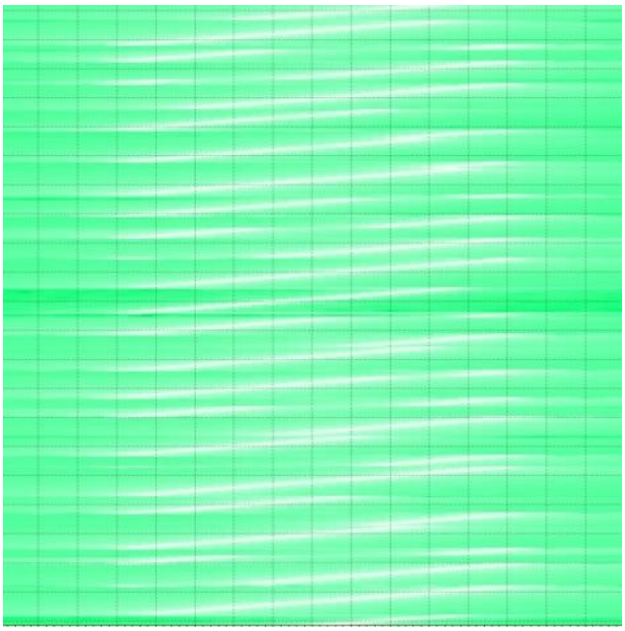


Рисунок 4. Частотно-временная панорама тестового сигнала с ЛЧМ модуляцией

.

Применение SDR-платформы с SoC на борту делает устройство обработки сигналов более компактным, а также расширяет диапазоны работы с различными системами связи. При проводимых ранее экспериментах, совмещение, например БСС на чипсетах NRF24L01+ (RF24Mesh) и модулях ESP32 (PainlessMesh) в ISM диапазоне, требовало наличия отдельного микроконтроллера для каждого чипсета, а малая вычислительная способность микроконтроллеров семейства Atmega32U4 затрудняла обработку данных непосредственно на самих узлах сети. Например, применение интегральных метрик для БСС [10] практически невозможно, ввиду ограниченной вычислительной мощности и малого объема памяти указанного семейства микроконтроллеров.

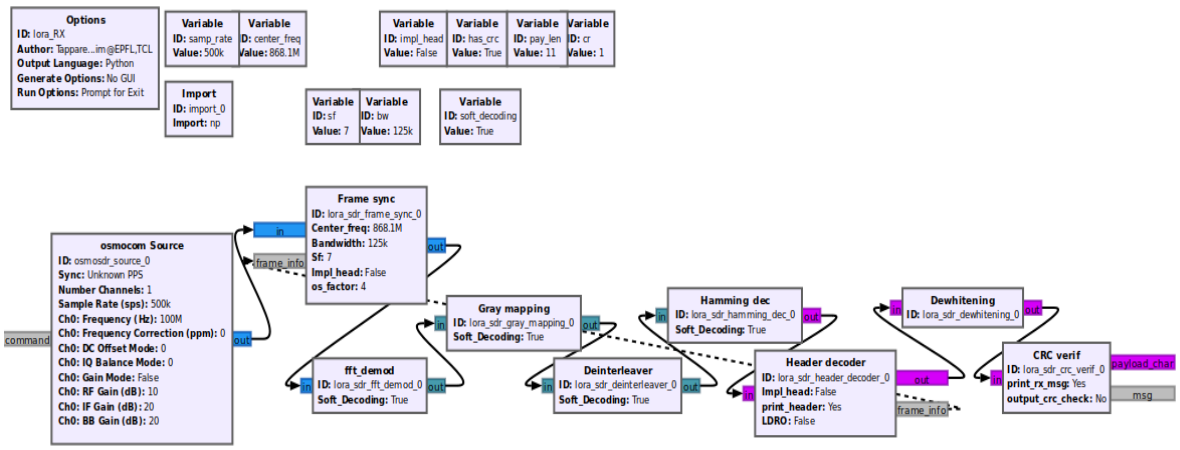


Рисунок 5. Схема тракта демодуляции и декодирования библиотеки gr-lora_sdr

Заключение

По результатам проведенных экспериментов можно сделать вывод, что применение SDR-платформ с SoC имеет высокий потенциал в области беспроводных самоорганизующихся сетей, в том числе и беспроводных сенсорных сетей, а также удобно в использовании в

качестве универсального радиоприемного устройства в мобильной робототехнике. Объединение одноплатного вычислительного устройства с широкодиапазонным АЦП/ЦАП позволяет получить мощный инструмент для проведения исследований и экспериментов в различных областях мобильной робототехники. Использование названных выше SDR-устройств накладывает определенные габаритные ограничения на мобильных роботов, несущих одноплатное вычислительное устройство на борту, однако наличие выводов GPIO в устройствах Pluto Plus позволяет управлять периферией автономной мобильной платформы напрямую, без микроконтроллера.

Список литературы

1. Галкин, В. А. Основы программно-конфигурируемого радио / В. А. Галкин; Мин-во образования и науки Российской Федерации, Нац. исследовательский ун-т, Московский гос. ин-т электронной техники (технический ун-т). М.: МИЭТ, 2010. 224 с.
2. Фокин, Г. А. Технологии программно-конфигурируемого радио / Г. А. Фокин. М.: Научно-техническое издательство "Горячая линия-Телеком", 2019. 316 с. – ISBN 978-5-9912-0819-2. – EDN TMJLHA.
3. Фокин, Г. А. Экспериментальная апробация SDR-платформы LibreSDR / Г. А. Фокин, К. Е. Рютин // Научно-техническая конференция Санкт-Петербургского НТО РЭС им. А.С. Попова, посвященная Дню радио. 2024. № 1(79). С. 174-177. – EDN CKWRVY.
4. Халимов, Н. Р. Распределенная сетецентрическая система управления группой ударных беспилотных летательных аппаратов / Н. Р. Халимов, А. В. Мефедов // Системы управления, связи и безопасности. 2019. № 3. С. 1–13. DOI: 10,24411/2410-9916-2019-10301
5. Библиотека RF24Mesh. URL: http://tmrh20.github.io/RF24Mesh/index.html (дата обращения 06.12.2024).
6. Библиотека Arduino LoRa. URL: https://github.com/sandeepmistry/arduino-LoRa/blob/master/API.md (дата обращения: 16.01.2025).
7. Sniffing and decoding NRF24L01+. URL: http://blog.cyberexplorer.me/2014/01/sniffing-and-decoding-nrf24l01-and.html (дата обращения 26.10.2024).
8. Библиотека gr-lora. URL: https://github.com/rpp0/gr-lora (дата обращения: 14.01.2025).
9. Библиотека gr-lora_sdr. URL: https://github.com/tapparelj/gr-lora_sdr (дата обращения: 14.01.2025).
10. Гурьянов, Е. Д. Применение интегральной метрики в многоскачковых беспроводных сенсорных сетях / Е. Д. Гурьянов // Проблемы информатики в образовании, управлении, экономике и технике: сборник статей XXIV Международной научно-технической конференции. Пенза, 22–23 ноября 2024 года. Пенза: Пензенский государственный университет, 2024. С. 184-189. – EDN ECBZRN.

APPLICATION OF A SOFTWARE DEFINED RADIO IN A CLUSTER WIRELESS SENSOR NETWORK

E.D. Guryanov

Demidov Yaroslavl State University

Annotation. An analysis of the prospects for using a software defined radio (SDR) as a receiving and transmitting equipment for simulating a node of a cluster wireless sensor network (WSN) for autonomous mobile platforms is presented. The study examines an algorithm for decoding received messages in an application to a multi-hop sensor network with a communication system based on NRF24L01+ chipsets (ESB protocol) and a radio network based on LoRa modules. The possibility of equipping autonomous mobile platforms with SDR receivers is being considered, as well as additional opportunities offered by the use of SDR in mobile robotics.

ВИЗУАЛИЗАЦИЯ ТРЕХМЕРНЫХ ДАННЫХ ПРИ ПОМОЩИ ГАУССОВА СПЛЭТИНГА

преп. Кирнос В.П., студ. Баев М.С., студ. Курносов Д.Г., студ. Максименко Е.Е., проф. Приоров А.Л.

Ярославский государственный университет им. П.Г. Демидова

Современные технологии рендеринга требуют эффективных методов визуализации сложных сцен. Одним из перспективных подходов является гауссов сплэтинг – метод, использующий трёхмерные гауссианы для представления объектов и сцен. В отличие от традиционных полигональных сеток, этот метод позволяет добиться высокой детализации при меньших вычислительных затратах. Гауссов сплэтинг строит сцену на основе облаков точек, преобразуемых в гауссианы, что обеспечивает плавные переходы и реалистичное освещение. Метод адаптивен: плотность гауссиан варьируется в зависимости от уровня детализации, что снижает нагрузку на память и процессор. Благодаря этим свойствам он находит применение в научной визуализации, 3D-сканировании, а также в компьютерных играх и VR/AR. Рассматриваются алгоритмы рендеринга и ключевые преимущества гауссова сплэтинга, а также его ограничения и перспективы развития. Одной из ключевых задач остаётся оптимизация метода для работы в реальном времени и интеграция с современными графическими движками.

Введение

Современные технологии рендеринга и визуализации трёхмерных данных постоянно развиваются, предлагая всё более совершенные инструменты для работы с цифровыми сценами. Одной из таких инноваций является гауссов сплэтинг – метод, использующий математическое представление в виде трёхмерных гауссиан для описания объектов и сцен [1-2]. Это открывает новые возможности в создании фотореалистичных изображений с минимальными вычислительными затратами, что является одной из важных задач систем технического зрения.

Ключевая идея гауссова сплэтинга заключается в замене традиционных подходов к описанию сцены, таких как использование полигональных сеток, на компактное представление в виде облаков точек с гауссианами [3-4]. Это позволяет повысить производительность, снизить потребление памяти и добиться высокой детализации. Такие свойства делают метод особенно перспективным для научных исследований, визуализации сложных данных, разработки компьютерных игр, а также систем виртуальной и дополненной реальности [5-6].

В работе рассматриваются основные аспекты гауссова сплэтинга, включая процесс рендеринга, а также имеющиеся преимущества и возможные ограничения. Кроме того, обсуждается его применение в различных областях и перспективы дальнейшего развития.

Основные принципы гауссова сплэтинга

Гауссов сплэтинг представляет трёхмерные сцены в виде совокупности гауссиан, каждый из которых описывает распределение плотности в пространстве. Этот подход кардинально отличается от традиционных методов, таких как использование полигональных сеток, которые требуют сложной геометрической структуры и текстур. Вместо этого гауссов сплэтинг опирается на следующие ключевые аспекты.

1. Представление сцены

В основе метода лежит использование гауссиан, которые представляют собой трёхмерные распределения плотности. Каждый гауссиан определяется набором параметров, включая центр, ковариационную матрицу, которая задаёт форму и ориентацию, а также цвет и прозрачность. Эти параметры позволяют гибко описывать как мелкие детали сцены, так и ее более крупные структуры.

Для создания сцены сначала собирается облако точек, которое затем преобразуется в набор гауссиан. Такой подход особенно полезен для обработки данных, полученных из реальных источников, например, от 3D-сканеров или из фотографий.

2. Рендеринг с использованием гауссиан

Процесс рендеринга в гауссовом сплэтинге заключается в проецировании гауссиан на плоскость экрана. Каждый гауссиан интерполируется с учётом его пространственного положения, размера и цветовых характеристик. Использование гауссиан позволяет создавать плавные переходы между различными элементами сцены, что делает изображение визуально приятным и реалистичным.

Для выполнения рендеринга могут использоваться как растеризация, так и трассировка лучей. В обоих случаях метод эффективно работает благодаря компактности представления данных. Пример построения представлен на рис.1. При этом возможна оптимизация вычислений путём адаптации уровня детализации.



а) Фотограмметрия б) Метод гауссова сплэтинга

Рисунок 1 – Примеры создания трехмерных сцен с использованием фотограмметрии и метода гауссова сплэтинга

3. Адаптивность и оптимизация

Одним из ключевых преимуществ гауссова сплэтинга является его адаптивность. В сценах с высокой детализацией можно использовать больше гауссиан, чтобы сохранить мелкие элементы. В то же время, менее значимые области могут быть представлены меньшим количеством гауссиан, что экономит вычислительные ресурсы и память.

Такая адаптивность делает метод особенно эффективным при работе с большими сценами, где важно сбалансировать качество изображения и необходимую производительность вычислительной системы.

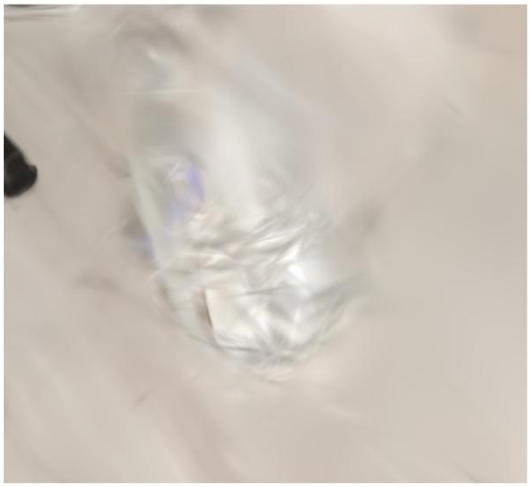
Преимущества и недостатки гауссова сплэтинга

Гауссов сплэтинг предлагает ряд уникальных преимуществ, которые делают его предпочтительным выбором в сравнении с традиционными методами рендеринга:

1. *Высокая производительность*. Использование гауссиан позволяет существенно снизить вычислительные затраты, особенно при обработке больших сцен.
2. *Экономия памяти.* Благодаря компактному представлению данных метод требует значительно меньше компьютерной памяти, чем традиционные полигональные или воксельные структуры.
3. *Плавность и фотореализм.* Гауссов сплэтинг обеспечивает плавные переходы между элементами сцены, что делает его подходящим для создания реалистичных изображений.

4. *Устойчивость к шуму.* Метод хорошо справляется с шумными данными, такими как облака точек из 3D-сканеров.

Из недостатков следует отметить самый весомый – проблемы при работе с прозрачными и отражающими поверхностями (рис. 2).



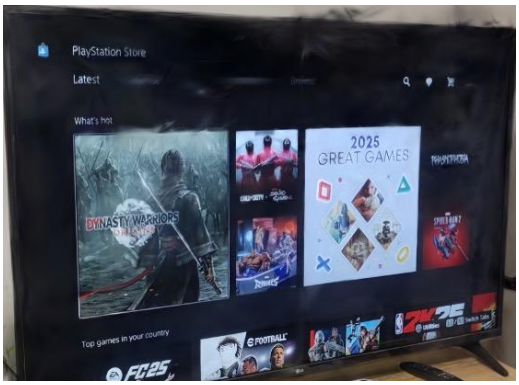
а) Обработанный прозрачный объект



б) Исходный прозрачный объект



в) Обработанная отражающая поверхность



г) Исходная отражающая поверхность

Рисунок 2 – Примеры создания трехмерных сцен с использованием разных методов

Области применения

Гауссов сплэтинг находит применение в различных областях благодаря своей универсальности и эффективности.

1. *Научная визуализация.* Гауссов сплэтинг используется для отображения сложных данных, таких как биологические структуры, астрономические объекты или модели молекул.
2. *Обработка 3D-сканов.* Метод идеально подходит для работы с облаками точек, полученными с помощью 3D-сканеров, и позволяет визуализировать их с высокой детализацией.
3. *Компьютерные игры.* Здесь метод позволяет рендерить сложные сцены в реальном времени, что особенно важно для игр с открытым миром.
4. *Виртуальная и дополненная реальности.* Высокая производительность и компактность данных делают метод подходящим для создания иммерсивных VR- и AR-опытов.

Сравнение с другими алгоритмами

Для построения трехмерных сцен используется не только алгоритм гауссова сплэтинга, но и классические в данной области алгоритмы: точечные облака, полигональный и воксельный методы. Основная цель применения указанных методов – изучение одновременной локализации и построения карты трехмерного пространства при движении в нем, то есть алгоритм SLAM с построением объемной карты окружающего пространства. Обычно при использовании SLAM нет возможности детального рассмотрения объектов со всех сторон, поэтому появляется ряд ограничений.

Главным вызовом здесь является отказ от датчиков глубины и использование монокамеры. Проведены эксперименты на основе PennCOSYVIO датасета, где для исследования использовалось 3-х минутное RGB-видео, снятое со скоростью 30 кадров в секунду (рис. 3).

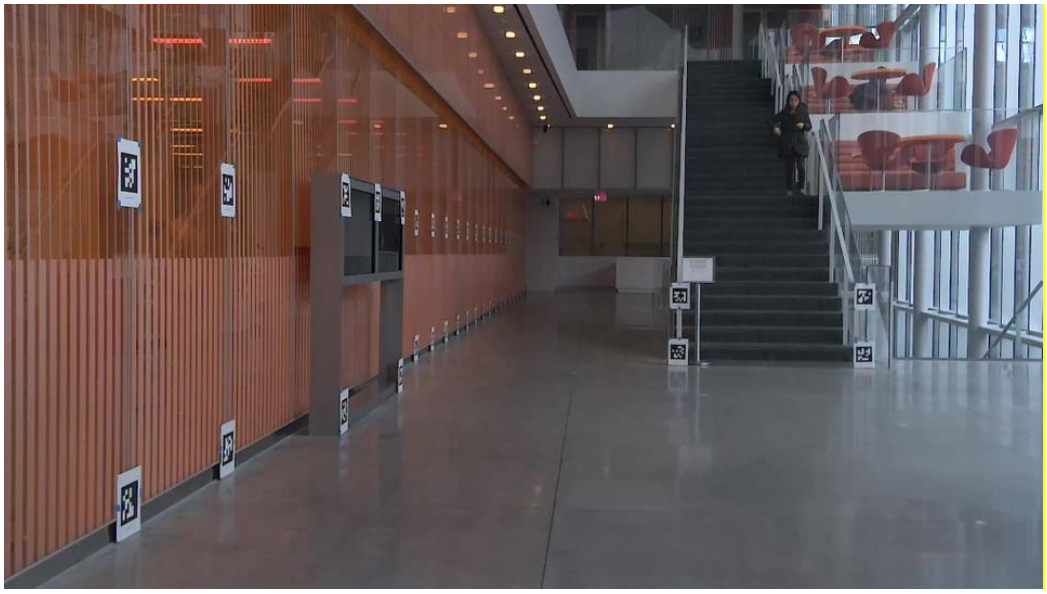


Рисунок 3 – Пример сцены из PennCOSYVIO датасета

Проведено сравнение точности, скорости и потребления памяти при использовании полигонального метода, воксельного метода и метода гауссова сплэтинга, результаты которого приведены в табл. 1.

Таблица 1. Сравнение методов построения трехмерного пространства по приведенным метрикам

Метод	Точность	Скорость	Память
Полигональный	9/10	6/10	5/10
Воксельный	7/10	8/10	7/10
Гауссов сплэтинг	10/10	5/10	6/10

Можно сделать вывод, что по скорости выигрывает воксельный метод, а по точности – гауссов сплэтинг.

Для отслеживания каждый из этих алгоритмов требует свой тип данных для корректной работы:

– Полигональный метод нуждается в плотных данных, когда почти каждая точка пространства имеет информацию о глубине и цвете. Так как используется алгоритм SLAM, то поиск всех точек пространства не выполним, а при расчетах используются только особые точки.

– Воксельный метод использует полуплотные данные. Здесь уже не требуется информация обо всех точках пространства, но основной вклад в формирование карты все равно вносят плотные облака.

– Гауссов сплэтинг позволяет создавать трехмерные карты на основе разреженных облаков точек. Это вариант лучше подходит к логике работы SLAM, когда нет уверенности, что особые точки будут формировать плотные облака.

Перспективы

Несмотря на многочисленные преимущества, гауссов сплэтинг сталкивается с рядом вызовов:

1. Оптимизация для реального времени. Обработка динамических сцен с изменяющимися параметрами требует дальнейшего совершенствования.
2. Интеграция с существующими инструментами. Важно адаптировать метод для использования в современных графических движках.
3. Масштабируемость. Работа с огромными объёмами данных требует оптимизации алгоритмов и улучшения их производительности.

В дальнейшем необходимо провести анализ представленного алгоритма с использованием систем бинокулярного зрения и трехмерных сканеров пространства.

Заключение

Гауссов сплэтинг представляет собой инновационный подход к рендерингу и визуализации трёхмерных данных. Его способность эффективно работать с облаками точек, обеспечивать фотореализм и экономить ресурсы делает его перспективным для множества приложений. Тем не менее, дальнейшее развитие метода, включая его оптимизацию и интеграцию, открывает ещё больше возможностей для использования в самых разных областях, в том числе научной визуализации и обработки данных, полученных с 3D-сканеров.

Литература

1. Kerbl B., Kopanas G., Leimkühle, T., Kautz J., & Nalbach O. (2023). 3D Gaussian Splatting for Real-Time Radiance Field Rendering // ACM Transactions on Graphics (TOG), 42(4), 1-12.
2. Müller T., Evans A., Schied C., & Keller A. (2022). Instant Neural Graphics Primitives with a Multiresolution Hash Encoding // ACM SIGGRAPH 2022 Conference Proceedings.
3. Zwicker M., Pfister H., Van Baar J., & Gross M. (2002) // EWA Splatting. IEEE Transactions on Visualization and Computer Graphics, 8(3), 223-238.
4. Rusinkiewicz S., Levoy M. (2000). QSplat: A Multiresolution Point Rendering System for Large Meshes // Proceedings of SIGGRAPH 2000, 343-352.
5. Lecun Y., Bengio Y., & Hinton G. (2015). Deep Learning // Nature, 521(7553), 436-444.
6. Pajarola R., & Gobbetti E. (2015). Survey of Semi-Structured Point-Based Rendering Methods // Computer Graphics Forum, 34(1), 178-201.

VISUALIZATION OF THREE-DIMENSIONAL DATA USING GAUSSIAN SPLATTING

Kirnos V.P., Baev M.S., Kurnosov D.G., Maksimenko E.E., Priorov A.L.

P.G. Demidov Yaroslavl State University

Modern rendering technologies require effective visualization methods for complex scenes. One of the promising approaches is Gaussian splitting, a method that uses three–dimensional Gaussians to represent objects and scenes. Unlike traditional polygonal meshes, this method allows you to achieve high detail with less computational effort. Gaussian splashing builds a scene based on clouds of points converted into Gaussians, which ensures smooth transitions and realistic lighting. The method is adaptive: the density of Gaussians varies depending on the level of detail, which reduces the load on memory and processor. Due to these properties, it finds application in scientific visualization, 3D scanning, as well as in computer games and VR/AR. Rendering algorithms and the key advantages of Gaussian splitting, as well as its limitations and development prospects are considered. Optimization of the method for real-time operation and integration with modern graphics engines remains one of the key tasks.

ИССЛЕДОВАНИЕ НЕЙРОСЕТЕВОЙ АРХИТЕКТУРЫ YOLO ДЛЯ ИДЕНТИФИКАЦИИ ЖЕЛЕЗНОДОРОЖНЫХ КОЛЕС В ВИДЕОПОТОКЕ С ИНКРЕМЕНТНЫМ ОБУЧЕНИЕМ

асп. Мареев А.В.[1], д.т.н. Орлов А.А.[1], студ. Верещагин И.Д.[1]

[1] *Владимирский государственный университет имени Александра Григорьевича и Николая Григорьевича Столетовых, г. Владимир, Российская Федерация*

В статье рассматривается задача обеспечения прослеживаемости продукции на линии производства железнодорожных колес при помощи компьютерного зрения. Цель работы — разработка и исследование алгоритма идентификации железнодорожных колес по их маркировке для обеспечения прослеживаемости продукции, что позволит повысить эффективность производства. Алгоритм идентификации железнодорожных колес основан на нейросетевых моделях YOLOv11 и включает следующие шаги: локализацию колеса, локализацию области маркировки, выравнивание и кадрирование полученной области, сегментацию и классификацию символов маркировки. На шаге классификации символов маркировки применяется инкрементный подход обучения с целью повышения точности распознавания в процессе эксплуатации систем, разработанных на основе предложенного алгоритма. В работе использованы следующие наборы данных: 250 изображений колес, 5000 изображений десяти классов цифр в градациях серого размером 28 на 28 пикселей. Приведены результаты исследования точности и скорости обработки изображений нейросетевыми моделями. Результаты исследований будут использованы для разработки системы прослеживания продукции в реальном времени.

Введение

В современных производственных системах контроль качества и прослеживаемость продукции имеют критическое значение из-за сложности процессов производства и логистики. В данной работе рассматривается производство такой продукции как железнодорожные колеса. Для обеспечения непрерывного мониторинга и эффективного производства необходимо идентифицировать железнодорожные колеса на каждом этапе производства. Данная задача является актуальной, так как существует множество работ по данному направлению в различных областях промышленности и сервиса [1-6]. Например, в работе [1] рассматривается задача распознавания текста с достаточно контрастных поверхностей, не имеющих значительных проективных искажений. Авторы работ [2,3] изучают более сложные задачи распознавания текстов с вывесок и автомобильных номеров, также имеющих контрастный текст, но уже с большими проективными искажениями. В работах [4,5] авторы исследуют проблему, напрямую касающуюся распознавания символов с металлических поверхностей. Но недостатком этих работ является малое допустимое проективное искажение и использование сложной конструкции освещения объекта интереса. Авторы работы [6] исследуют наиболее похожий на изучаемый в нашей работе объект интереса (распознавание маркировки тормозных дисков), но используется достаточно сложный метод на основе фотометрического стерео, что усложняет применение данного подхода к громоздким железнодорожным колесам. В большинстве рассмотренных работ используются глубокие нейронные сети. В настоящий момент исследования архитектуры такой нейронной сети как YOLO демонстрируют высокие показатели точности и скорости обработки среди других глубоких нейронных сетей для локализации объектов на изображении [7-14]. Последней версией архитектуры является YOLOv11, которая, в свою очередь, по результатам исследований, проведенных в работах [15,16], является оптимальным выбором для использования в системах прослеживания. Данная работа сосредоточена на разработке алгоритма, основанного на архитектуре нейронной сети YOLOv11, который автоматизирует процесс локализации и распознавания маркировки на железнодорожных колесах. Это не только повысит эффективность производства, но и обеспечит прослеживание продукции на всех этапах производственного цикла. Настоящая статья содержит материал по развитию системы распознавания маркировки железнодорожных колес, описанной в работах [17] и [18].

Несмотря на высокую точность классификации объектов на изображениях при помощи нейросети YOLO, обученной на статических данных большого объема, имеющих одинаковое распределение и хорошую разметку, в процессе эксплуатации модели накапливаются новые данные, которые в итоге неверно классифицируются. Традиционно для повышения точности в таких ситуациях модель нейросети заново обучают на всем объеме данных, затратив большое

количество времени без возможности сразу использовать полученные новые знания об объектах классификации. Поэтому современным моделям необходимо приспосабливаться к этим изменениям и обучаться непрерывно, сохраняя требуемый уровень точности результатов. Для решения данной проблемы используется инкрементный метод обучения, основанный на принципе «сна». Метод описан в работе [19, 20].

Описание набора данных

В соответствии с ГОСТ 10791–2011 «КОЛЕСА ЦЕЛЬНОКАТАНЫЕ» на железнодорожное колесо наносят маркировку в следующем порядке: две последние цифры года изготовления колеса, марка стали колеса, номер плавки колеса, условное обозначение предприятия (изготовителя колеса), порядковый номер колеса по системе нумерации предприятия-изготовителя. Между условным обозначением предприятия-изготовителя и порядковым номером колеса предусмотрено место шириной от 30 до 150 мм для приемочных клейм заказчика. Маркировку наносят в горячем состоянии на боковую поверхность обода с наружной стороны колеса. Высота знаков маркировки должна составлять от 10 до 12 мм, глубина знаков - до 4 мм. Расстояние от основания знаков маркировки до кромки внутренней поверхности обода с наружной стороны колеса составляет от 3 до 10 мм. При нанесении маркировки используются цифры, буквы и другие знаки маркировки.

В качестве входных данных разрабатываемого алгоритма идентификации колес используется два набора данных. Первый набор данных включает в себя 250 изображений колес размером 2600 на 2110 пикселей. Второй набор данных содержит изображения цифр в градациях серого, распределенные по 10 классам и общей численностью 5000. Размер каждого изображения масштабирован к 28 на 28 пикселей. Пример изображения колеса с нанесенной маркировкой представлен на рисунке 1. Сведения об объемах наборов данных для обучения приведены в таблице 1 и 2.



Рисунок 1 – Пример изображения маркировки на железнодорожном колесе: 1 - две последние цифры года изготовления колеса; 2 - марка стали; 3 - номер плавки; 4 - условный номер предприятия-изготовителя колеса (номер или его торговая марка); 5 - порядковый номер колеса по системе нумерации предприятия-изготовителя

Таблица 1. Сведения об объемах набора изображений железнодорожных колес

Название выборки	**Количество изображений**	**Процентное соотношение**
Тренировочная выборка	160	64%
Валидационная выборка	40	16%
Тестовая выборка	50	20%

Таблица 2. Сведения об объемах набора изображений символов маркировки

Название выборки	**Количество изображений**	**Процентное соотношение**
Тренировочная выборка	2800	56%
Валидационная выборка	700	14%
Тестовая выборка	500	10%
Выборка для инкрементного обучения	1000	20%

Алгоритм идентификации железнодорожных колес с помощью нейросети YOLO

В настоящее время на предприятии используется стационарная система компьютерного зрения, требующая специальной подсветки и точного позиционирования колеса относительно ракурса камеры. В отличие от этого предложенный в данной работе алгоритм обладает рядом преимуществ: системе, разработанной на основе этого алгоритма, не потребуется точное позиционирование колеса, нет необходимости обеспечивать идеальное освещение, а также ее можно будет устанавливать на других участках производства без значительных затрат, так как

система имеет малые габариты. Кроме того, данный алгоритм позволит создать мобильное приложение для идентификации колес на складах, в вагонных и локомотивных депо, что делает его более гибким и универсальным решением для производственных и эксплуатационных задач.

Локализация маркировки в видеопотоке затруднена наличием помех, таких как недостаточная освещенность объекта интереса (маркировки) или колеса, блики, перекрытие объекта интереса другим, выход объекта интереса из области съемки. Учитывая данные особенности, предлагается реализовать локализацию маркировки железнодорожного колеса с помощью нейросети YOLO [21] на основе алгоритма, состоящего из следующих шагов:

Шаг 1. Детекция колес и их трекинг. Из видеопотока извлекается текущий кадр, затем с помощью нейросети YOLOv11, обученной на детекцию колес, выполняется обнаружение колес на изображении с присвоением им уникального индекса. Полученный индекс передается в следующие шаги вместе с изображением этих колес. Если имеются объекты с известным индексом и присвоенной маркировкой на шаге 7, то изображения этих колес передаются на 2 шаг каждые 10 обнаружений для проверки корректности трекинга. Если колес не обнаружено, то перейти к началу 1 шага.

Шаг 2. Локализация области маркировки. Для выделения области маркировки используется нейросеть YOLOv11-obb, обученная на базе данных с размеченной маркировкой. Эта модель определяет область, в которой расположена маркировка, с учетом различных положений и условий освещения. Если маркировку не удалось локализовать, то переход к 1 шагу.

Шаг 3. Выравнивание и кадрирование области маркировки. После обнаружения маркировки выполняется поворот и кадрирование области с учетом параметров ориентации, что упрощает дальнейшее распознавание символов. Угол, на который нужно повернуть изображение, определяется исходя из положения колеса и маркировки.

Шаг 4. Сегментация и обнаружение символов. На данном шаге нейросеть YOLOv11-seg, обученная на сегментации символов и брака маркировки, локализует все объекты интереса.

Шаг 5. Распознавание символов. Нейросеть YOLOv11-cls классифицирует обнаруженные символы. Если при классификации не удалось преодолеть порог уверенности в 70% больше, чем у 10 символов, то символы, которые были классифицированы с уверенностью больше заданного порога, сохраняются вместе с индексом колеса, присвоенном ему на шаге 1, и выполняется переход к шагу 1. Если количество повторных итераций классификации символов для текущего индекса объекта превышает 20 без достижения порога уверенности в 90%, система уведомляет оператора о необходимости вмешательства для ручной классификации обнаруженных символов.

Шаг 6. Оценка уверенности и контроль качества. Если количество символов, классифицированных с уверенностью 95% и выше, превышает 16 (минимальное количество символов в маркировке), то символы собираются в строку и записываются в базу данных системы прослеживания вместе с индексом объекта из 1 шага. Переход к шагу 7. Если же данный порог не пройден, то символы, у которых уверенность ниже 95% добавляются в базу данных для инкрементного обучения и отображаются для оператора с целью их классификации.

Шаг 7. Оператору отображается индекс объекта и текст распознанной маркировки. Если количество символов для инкрементного обучения больше 200, то оператор уведомляется о необходимости итерации инкрементного обучения. Поскольку производство партии железнодорожных колес нельзя остановить, оператор выбирает удобное время для «сна» системы, чтобы дообучить модель YOLOv11-cls. Переход к шагу 1.

Результаты исследования

Для шага 2 локализации области маркировки в видеопотоке из алгоритма проведено исследование с целью выбора оптимальной модели нейросети YOLO. Для этого было проведено сравнение различных версий модели по скорости работы (FPS — количество кадров в секунду) и точности (mAP50 и mAP75 — средняя точность на уровнях 50% и 75%). Результаты исследований приведены в таблице 3.

Таблица 3. Сравнение моделей YOLO

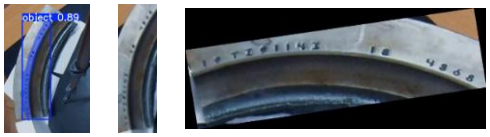
Модель	FPS	mAP50	mAP75	max FPS	min FPS
yolov11n.pt	9,7	0,607	0,476	10,78	7,2
yolov11s.pt	3	0,760	0,638	3,1	2,73
yolov11m.pt	1,2	0,784	0,652	1,53	0,8
yolov11l.pt	0,6	0,828	0,706	0,61	0,59

Наибольший интерес в исследовании был к модели YOLOv11s, учитывая ее оптимальное соотношение точности (mAP50, mAP75) 0,760 и 0,638, а также скорости работы. Далее было проведено исследование подбора количества эпох для обучения нейросети YOLOv11s, что позволит определить оптимальные настройки для достижения удовлетворительного качества распознавания. Результаты приведены в таблице 4. Оптимальные результаты были достигнуты при количестве эпох, равном 200.

Таблица 4. Экспериментальные результаты для выбора количества эпох обучения

Количество эпох	Потеря для рамок	mAP50	mAP50-95
50	0,824	0,993	0,805
100	0,668	0,995	0,890
200	0,368	0,995	0,938
300	0,341	0,995	0,921

На основе полученных результатов была дообучена нейронная сеть YOLOv11s, которая успешно обнаруживает маркировку на железнодорожном колесе на шаге 2. Пример обнаружения на основе предложенного алгоритма представлен на рисунке 2-а.



(а) (б) (в)

Рисунок 2 – Изображение с обнаруженной разметкой

Затем из полученных данных была выделена область, содержащая маркировку. Вырезанный прямоугольный фрейм, полученный от нейросети, показан на рисунке 2-б.

Далее, исходя из известных координат и угла поворота, изображение маркировки выравнивается и кадрируется. Пример поворота изображения демонстрируется на рисунке 2-в.

Оценка точности обученных моделей, используемых на шагах 2 и 5, представлена в таблице 5.

Таблица 5. Оценка точности

Модель	Метрика	Описание метрики	Значение
YOLOv11-obb	Precision	Точность локализации объектов.	0,945
	Recall	Полнота локализации объектов.	0,995
	mAP50	Средняя точность при пороге IoU 50% для локализации объектов.	0,982
	mAP50-95	Средняя точность при порогах IoU от 50% до 95% для локализации объектов.	0,932
YOLOv11-cls	Precision	Точность распознавания символов.	0,957
	Recall	Полнота распознавания символов.	0,942
	mAP50	Средняя точность при пороге IoU 50% для распознавания символов.	0,961
	mAP50-95	Средняя точность при порогах IoU от 50% до 95% для распознавания символов.	0,657

С целью исследования применения инкрементного обучения было проведено три итерации дообучения. На рисунке 3 представлена матрица ошибок после инкрементного дообучения.

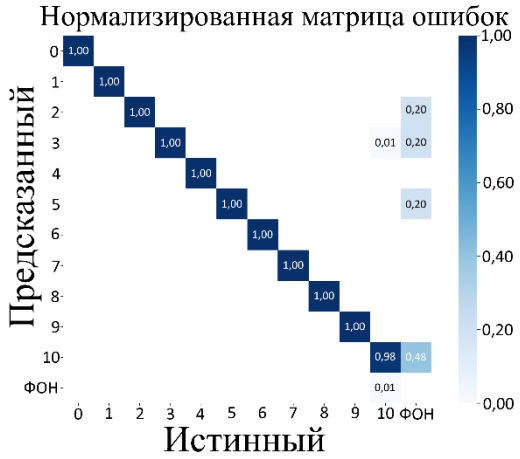


Рисунок 3 – Матрица ошибок после инкрементного дообучения

Заключение

В ходе проведенных исследований были достигнуты следующие результаты. Среди моделей YOLO оптимальной по соотношению точности и производительности является модель YOLOv11s, которая обеспечивает высокие показатели точности (mAP50, mAP75) 0,760 и 0,638 при приемлемой скорости обработки. Для достижения данной точности потребовалось выполнить 200 эпох обучения. Обученная модель YOLO успешно обнаруживает маркировку на железнодорожных колесах и корректно выделяет нужную область с оценкой полноты локализации объектов, равной 0,995. Распознавание символов успешно выполняется на основе YOLOv11-cls с возможностью инкрементного дообучения.

Исследование выполнено при поддержке гранта Российского научного фонда, номер гранта 24-21-00386. Более подробную информацию о гранте можно найти по ссылке: https://rscf.ru/project/24-21-00386/.

Литература

1. Kralicek J., Matas J., Busta M. Care Label Recognition // 2019 International Conference on Document Analysis and Recognition (ICDAR). Sydney, NSW, Australia, 2019. pp. 959-966. doi: 10.1109/ICDAR.2019.00158.
2. Shekar Goud D., Vigneshwari M., Aparna P., Vijayasekaran G., Singh Yadav A., Kumar A. Text Localization and Recognition from Natural Scene Images using AI // 2022 International Conference on Automation, Computing and Renewable Systems (ICACRS). Pudukkottai, India, 2022. pp. 1153-1158. doi: 10.1109/ICACRS55517.2022.10029220.
3. Tourani A., Shahbahrami A., Soroori S., Khazaee S., Suen C.Y. A Robust Deep Learning Approach for Automatic Iranian Vehicle License Plate Detection and Recognition for Surveillance Systems // IEEE Access. 2020. Vol. 8. pp. 201317-201330. doi: 10.1109/ACCESS.2020.3035992.
4. Xiang Z., Wu H., Zhou D. Recognition of characters on curved metal workpiece surfaces based on multi-exposure image fusion and deep neural networks // Neurocomputing. 2022. Vol. 512. pp. 117-129. ISSN 0925-2312. doi: 10.1016/j.neucom.2022.09.022.
5. Xiang Z., Huang Y. A Metal Character Enhancement Method based on Conditional Generative Adversarial Networks // 2022 International Conference on Machine Vision, Automatic Identification and Detection (MVAID 2022). 2022. Vol. 2284. doi: 10.1088/1742-6596/2284/1/012003.
6. Liu S., Li Y., Zhao A., Liu K., Wang J. A brake disc character recognition method based on photometric stereo and surface curvature extraction // 2019 Chinese Automation Congress (CAC). Hangzhou, China, 2019. pp. 3324-3329. doi: 10.1109/CAC48633.2019.8996247.
7. Srivastava S. Comparative analysis of deep learning image detection algorithms // Journal of Big Data. 2021. Vol. 8. № 66. doi: 10.1186/s40537-021-00434-w.
8. Ерохин Д.Ю., Ершов М.Д. Современные сверточные нейронные сети для обнаружения и распознавания объектов // Цифровая обработка сигналов. 2018. № 3. С. 64-69.

9. Котов Н.В., Хрящев В.В., Тихомиров А.А. Исследование нейросетевой архитектуры YOLO для обнаружения изображения купола слепой кишки в колоноскопии // Цифровая обработка сигналов и ее применение DSPA – 2024. Москва, 2024. С. 151-155.
10. Жгутов П.Е., Муравьев В.С. Сравнительное исследование нейросетевых алгоритмов в задаче оценки плотности автомобильного потока // Цифровая обработка сигналов и ее применение DSPA – 2023. Москва, 2023. С. 205-209.
11. Кирнос В.П., Коковкина В.А. Разработка системы одновременной локализации и картографии на базе нейронных сетей // Цифровая обработка сигналов и ее применение DSPA – 2023. Москва, 2023. С. 263-266.
12. Потапов Е.В., Артамонов К.С., Сидорова Е.И. Сверточная нейронная сеть YOLOv7 для детектирования объектов интереса на изображениях со сферического робота // Цифровая обработка сигналов и ее применение DSPA – 2023. Москва, 2023. С. 367-371.
13. Васильев Д.А., Левкин Т.В., Сконников П.Н., Трофимов Д.В. Модификация детектора объектов YOLO для реализации на ПЛИС в реальном времени // Цифровая обработка сигналов. 2022. № 3. С. 77-79.
14. Хрящев В.В., Приоров А.Л., Лебедев А.А., Котов Н.А. Обнаружение полипов на колоноскопических изображениях при помощи алгоритмов на базе нейросетевой архитектуры YOLO // Цифровая обработка сигналов. 2023. № 3. С. 27-31.
15. Jegham N., Koh C.Y., Abdelatti M., Hendawi A. Evaluating the Evolution of YOLO (You Only Look Once) Models: A Comprehensive Benchmark Study of YOLO11 and Its Predecessors // 2024. doi: 10.48550/arXiv.2411.00201.
16. Sapkota R., Karkee M. Comparing YOLO11 and YOLOv8 for instance segmentation of occluded and non-occluded immature green fruits in complex orchard environment // 2024. doi: 10.48550/arXiv.2410.19869.
17. Mareev A., Orlov A., Rybkin R. Localization of Railway Wheel Markings Based on Preliminary Localization of Symbols // 2023 DSPA. Moscow, Russian Federation, 2023. pp. 1-4. doi: 10.1109/DSPA57594.2023.10113361.
18. Mareev A., Orlov A. Development of a System for Localizing the Markings of Railway Wheels in a Video Stream // 2022 DSPA. Moscow, Russian Federation, 2022. pp. 1-4. doi: 10.1109/DSPA53304.2022.9790785.
19. Орлов А.А., Абрамова Е.С. Разработка и исследование алгоритма посменного инкрементного обучения нейронной сети // Компьютерная оптика. 2023. Т. 47, № 3. С. 491-498. doi: 10.18287/2412-6179-CO-1203.
20. Mareev A., Orlov A. Incremental Training of Convolutional Neural Networks for Recognition of Marking Symbols // 2024 ICIEAM. 2024. pp. 934-938.
21. YOLO Docs // Ultralytics. URL: https://docs.ultralytics.com (дата обращения: 26.12.2024).

INVESTIGATION OF YOLO NEURAL NETWORK ARCHITECTURE FOR IDENTIFICATION OF RAILWAY WHEELS IN VIDEO STREAM WITH INCREMENTAL LEARNING

Mareev A.V.[1], Orlov A.A.[1], Vereshchagin I.D.[1]

[1] *Vladimir State University named after Alexander and Nicolay Stoletovs, Vladimir, Russia*

The article discusses the problem of ensuring product traceability on the railway track production line using machine vision. The purpose of the work is to develop and study an algorithm for identifying railway wheels by marking them to ensure product traceability, which will improve production efficiency. The railway wheel identification algorithm is based on YOLOv11 neural network models and includes the following steps: wheel localization, localization of the marking area, alignment and cropping of the resulting area, segmentation, and classification of marking symbols. For the classification stage of marking symbols, an incremental learning approach is used to improve recognition accuracy during the operation of systems developed based on this algorithm. The following data sets were used in the work: 250 images of wheels, 5000 images of 10 classes of digits in grayscale measuring 28 by 28 pixels. The results of a study of the accuracy and speed of image processing by neural network models are presented. The research results will be used to develop a real-time product tracking system.

СРАВНИТЕЛЬНАЯ ОЦЕНКА МЕТОДОВ КОМПЕНСАЦИИ МЕТЕОРОЛОГИЧЕСКИХ ЭФФЕКТОВ НА ИЗОБРАЖЕНИЯХ ДИКИХ ЖИВОТНЫХ

асп. Наталенко Д.Н.

Федеральное государственное бюджетное образовательное учреждение высшего образования «Сибирский государственный университет науки и технологий имени академика М.Ф. Решетнева»

В докладе рассматриваются методы компенсации метеорологических эффектов, таких как дождь, снегопад и туман, на изображениях диких животных. Учитывая важность точности визуальных данных для экологии и охраны природы, представлены традиционные методы обработки изображений и современные нейросетевые подходы. Оценка их эффективности основана на экспериментальных данных и метриках качества изображений, что позволяет выявить сильные и слабые стороны каждого подхода. Результаты исследования могут быть полезны для специалистов в области компьютерного зрения, экологии и охраны дикой природы, а также для разработчиков программного обеспечения, занимающихся обработкой изображений. В докладе подчеркивается важность выбора оптимального метода для улучшения анализа изображений диких животных.

Введение

Компенсация метеорологических эффектов, таких как дождь, снегопад и туман, на изображениях диких животных существенно влияет на качество визуальных данных [1]. В свою очередь, это ограничивает эффективность последующего анализа, проводимого как вручную специалистами, так и автоматически, включая алгоритмы обнаружения и классификации. Из-за условий съемки на естественных территориях, когда разные метеорологические эффекты могут существовать как отдельно, так и совместно, задача их компенсации представляет собой сложную научную задачу и требует разработки специальных алгоритмов.

Современные достижения в области обработки изображений и глубокого обучения открывают новые возможности для решения данной проблемы. Исследования, рассматривающие традиционные методы компенсации метеорологических эффектов, выделяют различные подходы, такие как фильтрация по частотным компонентам, разделение частот и использование пространственно-временных фильтров. Хотя эти методы могут быть эффективны в определенных условиях, они часто не справляются с задачей восстановления изображения с комплексными метеорологическими явлениями. При этом исследования, использующие подходы с применением глубокого обучения, сконцентрированы на разработке нейросетевых моделей, способных адаптироваться к изменяющимся условиям съемки и обеспечивать более высокую точность восстановления изображений.

Каждый из предложенных подходов имеет свои преимущества и недостатки, которые необходимо выявить благодаря сравнительной оценке. В данной работе подчеркивается важность точного представления признаков для достижения высоких результатов в восстановлении изображений диких животных, ухудшенных погодными условиями. Представлен сравнительный анализ традиционных и нейросетевых методов компенсации метеорологических эффектов, акцентируя внимание на преимуществах и недостатках каждого подхода. Результаты исследования помогут определить оптимальные стратегии для повышения точности и скорости обработки визуальных данных в экологии, а также подчеркнут значимость междисциплинарного подхода в решении актуальных задач.

Традиционные и нейросетевые методы компенсации метеорологических эффектов

Многие современные методы компенсации метеорологических эффектов направлены на решение одной конкретной задачи, например, компенсации только дождя или только снегопада. При этом существуют и исключения, способные работать с комплексными задачами компенсации. Классификация рассматриваемых методов компенсации по типу и способам компенсации представлена на рисунке 1.



Рисунок 1. Классификация методов компенсации метеорологических эффектов

Среди рассматриваемых методов компенсации можно выделить следующие особенности:

– некоторые традиционные методы основаны на моделях физики атмосферы, что позволяет учитывать реальные метеорологические процессы. Это может быть полезно для задач, где важно учитывать физические законы;

– традиционные методы часто используют различные фильтры для устранения шумов и искажений, вызванных метеорологическими эффектами. Эти фильтры могут быть очень эффективными для определенных типов шумов и являются хорошо понятными и интерпретируемыми;

– традиционные методы часто требуют меньше параметров для настройки, что делает их более доступными для пользователей, не обладающих глубокими знаниями в области машинного обучения;

– нейросетевые методы могут не только анализировать данные, но и улучшать их качество. Например, с помощью генеративных моделей, которые могут восстанавливать отсутствующие или искаженные данные;

– нейросетевые методы основаны на обучении моделей и требуют большого количества данных, но при этом являются наиболее универсальными и точными в рамках работы со сложными данными;

– традиционные методы труднее адаптировать для работы сразу с несколькими задачами. Например, устранении сразу всех метеорологических эффектов сразу (дождь, снегопад и туман).

Среди рассматриваемых традиционных методов можно выделить подход с использованием TVG. Метод использует многонаправленный подход, что позволяет более точно обрабатывать различные направления дождевых капель. В свою очередь, использование неограниченного свертывания и колебаний позволяет эффективно сохранять детали изображения при удалении шумов. Среди рассматриваемых нейросетевых методов можно выделить подход с использованием простого межиерархического трансформера. Трансформеры хорошо справляются с задачами, которые связаны с вниманием, что особенно важно при наличии

шумов, таких как дождь и снег. Также метод использует кросс-иерархическую выборку, что позволяет акцентировать внимание на различных уровнях информации, улучшая качество восстановления изображений в сложных погодных условиях.

Если сравнивать подходы напрямую, то нейросетевые методы, как правило, превосходят традиционные методы в задачах, которые требуют высокой точности и адаптивности к различным условиям. Так, например, трансформеры и рекуррентные сети, могут достигать высокой точности в задачах удаления метеорологических эффектов благодаря способности обучаться на больших объемах данных и выявлять сложные паттерны. При этом, традиционные методы могут быть более эффективными в условиях ограниченных вычислительных ресурсов, поскольку нейросети требуют значительных затрат.

Сравнительный анализ методов компенсации метеорологических эффектов

В таблице 1 представлено сравнение оценок качества обработанных изображений, произведенное с помощью метрик: PSNR (Peak Signal-to-Noise Ratio), SSIM (Structural Similarity Index Measure), а также NIQE (Natural Image Quality Evaluator) для некоторых методов.

Таблица 1. Сравнение результатов тестирования

Метод компенсации	Категория эффектов	Набор данных	Результаты тестирования		
			PSNR, дБ	SSIM	NIQE
Удаление дождя и снега с помощью простого межиерархического трансформера	Дождь; Снегопад	RainCityscapes	35,16	0,99	–
		Snow100K-L	30,48	0,95	–
		Snow100K-M	34,44	0,97	–
		Snow100K-S	36,17	0,98	–
		RSC и RS100K	–	–	4,90
Многонаправленное удаление дождевых полос с использованием TGV	Дождь	–	24,76	0,89	–
		–	22,95	0,69	–
		–	20,60	0,80	–
Удаление дождя с помощью рекуррентных сетей с масштабным управлением	Дождь	Rain100H	28,98	0,90	–
		Rain100L	38,78	0,99	–
		Rain1200	32,93	0,93	–
Удаление дождевых полос с помощью адаптивной модели глубокого обучения со взвешенным сверточным словарем	Дождь	Rain100H	32,01	0,92	–
		Rain100L	40,63	0,99	–
		Rain1400	33,37	0,95	–
Удаление снега с помощью адаптивной остаточной сети с управлением маской снега	Снегопад	Snow100K-L	29,44	0,93	–
		Snow100K-M	33,22	0,95	–
		Snow100K-S	34,46	0,95	–
Удаление снега с помощью эффективной настройка параметров	Снегопад	CSD	39,16	0,98	–
		SRRS	34,13	0,97	–
		Snow100K	35,68	0,95	–
Удаление снега с помощью гибридной волновой трансформации	Снегопад	CSD	38,71	0,98	–
		SRRS	32,84	0,97	–
		Snow100K	35,37	0,94	–
Удаление тумана с помощью пирамиды ограничений норм	Туман	SOTS-Outdoor	25,64	0,87	–
Удаление тумана с помощью сети вейвлетов на основе трансформеров	Туман	O-Haze	21,32	0,72	–
		I-Haze	18,64	0,82	–
		NH-Haze	21,68	0,82	–
		RTTS	–	–	4,76
Устранение тумана с помощью улучшенной генерации	Туман	SOTS-Outdoor	28,15	0,96	–
		SOTS-Indoor	27,50	0,95	–
		Haze4k	26,12	0,95	–

Метрика PSNR измеряет отношение между максимальной возможной мощностью сигнала изображения и мощностью шума, который искажает это изображение. Метрика SSIM оценивает качество изображения, основываясь на восприятии структуры, яркости и контраста,

что лучше соответствует восприятию человеческим зрением. Метрика NIQE основана на статистических характеристиках естественных изображений и предназначена для оценки их качества с точки зрения восприятия.

Исходя из результатов, можно сделать вывод о следующем:

– нейросетевые методы удаления дождя и снега, такие как простой межиерархический трансформер, показывают наилучшие результаты. Это подтверждается высокими значениями PSNR и SSIM (например, PSNR=36,17 дБ и SSIM=0,98 для Snow100K-S), указывающими на их высокую эффективность в условиях сложной метеорологии;

– рекуррентные сети также демонстрируют отличные результаты, особенно в наборе данных Rain100L (PSNR=38,78 дБ, SSIM=0,99), что подчеркивает их способность к адаптации и точности в задачах удаления дождя;

– традиционные методы, такие как TGV, имеют значительно более низкие показатели качества (максимум PSNR=24,76 дБ, SSIM=0,89), что свидетельствует об их ограниченной эффективности по сравнению с современными нейросетевыми подходами;

– методы удаления снега, основанные на адаптивных остаточных сетях и эффективной настройке параметров, показывают конкурентоспособные результаты (например, PSNR=39,16 дБ и SSIM=0,98), что делает их одними из лучших в своей категории;

– удаление тумана демонстрирует значительно более низкие результаты по сравнению с другими эффектами, с наилучшим показателем PSNR=28,15 дБ. Это подчеркивает необходимость дальнейших исследований и улучшений в этой области, чтобы достичь сопоставимых с другими эффектами показателей качества.

Общая тенденция показывает, что выбор метода компенсации метеорологических эффектов должен основываться на конкретных условиях, доступных ресурсах и требованиях к качеству обработки изображений, что делает нейросетевые подходы предпочтительными для решения сложных задач.

Следующим этапом сравнения является визуальный анализ. Как правило, методы компенсации тестируются на разных наборах данных, что затрудняет их оценку. Тем не менее, визуальный анализ является важной частью сравнения методов. Примеры визуального сравнения работы методов компенсации представлены на рисунках 2-4.



а) б) в) г)

Рисунок 2. Визуальное сравнение результатов компенсации дождя:
а) TVG [2]; б) простой межиерархический трансформер [3]; в) рекуррентные сети [4]; г) модель со взвешенным сверточным словарем [5]



Рисунок 3. Визуальное сравнение результатов компенсации снегопада:
а) простой межиерархический трансформер [3]; б) адаптивной остаточная сеть [6]; в) эффективная настройка параметров [7]; г) гибридная волновая трансформация [8]

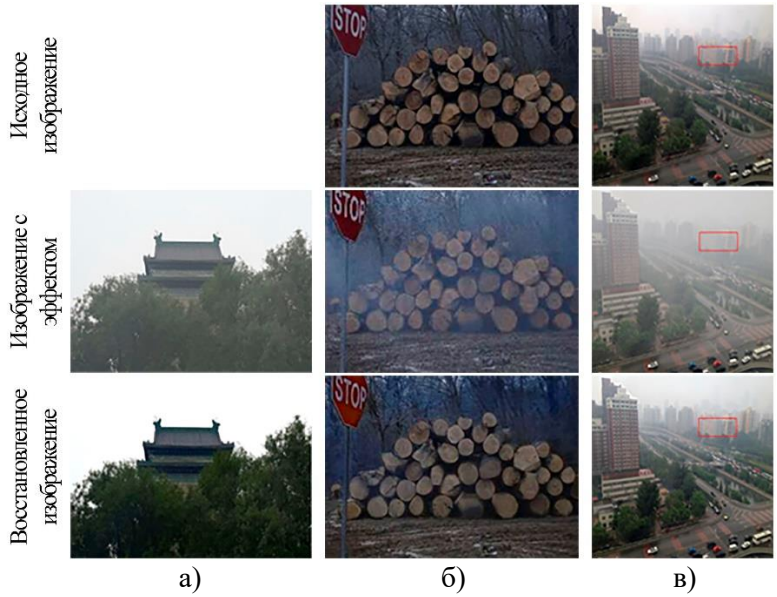


Рисунок 4. Визуальное сравнение результатов компенсации тумана:
а) пирамида ограничений норм [9]; б) сеть вейвлетов на основе трансформеров [10]; в) улучшенная генерация [11]

Исходя из результатов визуальной оценки, можно сделать вывод о неплохой работе традиционных методов компенсации дождя и тумана. Тем не менее, нейросетевые методы оставляют меньше артефактов после обработки и показывают лучшие результаты в случаях слишком обильных эффектов.

Заключение

Таким образом, проведенное сравнение подтвердило, что современные нейросетевые методы компенсации метеорологических эффектов на изображениях диких животных показывают значительно лучшие результаты по сравнению с традиционными подходами. Высокие значения PSNR и SSIM, полученные с использованием нейросетевых моделей, свидетельствуют о их способности эффективно восстанавливать изображения, ухудшенные дождем, снегом и туманом. Тем не менее, традиционные методы все еще могут быть актуальны

в определенных условиях, особенно когда ресурсы ограничены. Результаты сравнения могут служить основой для дальнейших разработок в области компьютерного зрения и обработки изображений.

Литература

1. О.Е. Гончарова. Влияние метеорологических условий видимости на характеристики формируемого изображения // Актуальные научные исследования в современном мире. 2020. № 12-2. С. 44-49.

2. Y. Gu, Y. Gao, H. Liu, D. Wang. Multi-directional rain streak removal based on infimal convolution of oscillation TGV // Neurocomputing. 2022. Vol. 486. P. 61-76.

3. Y. Wen, T. Gao, K. Zhang, P. Cheng, T. Chen. Restoring vision in rain-by-snow weather with simple attention-based sampling cross-hierarchy Transformer // Pattern Recognition. 2024. Vol. 156. Art. No. 110743.

4. C. Wang, H. Zhu, W. Fan, X.-M. Wu, J. Chen. Single image rain removal using recurrent scale-guide networks // Neurocomputing. 2022. Vol. 467. P. 242–255.

5. Y.-N. Zhang, L. Shen, Z. Lai. Adaptive weighted rain streaks model-driven deep network for single image deraining // Expert Systems with Applications: An International Journal. 2023. Vol. 222. Art. No. 119807.

6. B. Cheng, J. Li, Y. Chen, S. Zhang, T. Zeng. Snow Mask Guided Adaptive Residual Network for Image Snow Removal // Computer Vision and Image Understanding. 2022. Vol. 236. Art. No. 103819.

7. X. Dai, Y. Zhou, X. Qiu, H. Tang, T. Tong. Parameter-efficient fine-tuning for single image snow removal // Expert Systems with Applications: An International Journal. 2025. Vol. 265. Art. No. 125901.

8. X. Dai, Y. Zhou, X. Qiu, H. Tang, T. Tan, Q. Zhang, T. Tong. WaveFrSnow: Comprehensive perception wavelet transform frequency separation transformer for image snow removal // Digit. Signal Process. 2024. Vol. 155. Art. No. 104715.

9. L. Wu, H. Hu, G. Teng, Y. Yang, H. Zhang. Norm constraints pyramid for image dehazing // Digital Signal Processing. 2025. Vol. 156 (2). Art. No. 104828.

10. S. Zhang, Z. Tao, S. Lin. WaveletFormerNet: A Transformer-based wavelet network for real-world non-homogeneous and dense fog removal // Image and Vision Computing. 2024. Vol. 146. Art. No. 105014.

11. W. Li, D. Fan, Q. Zhu, Z. Gao, H. Sun. HEDehazeNet: Unpaired image dehazing via enhanced haze generation // Image and Vision Computing. 2024. Vol. 150. Art. No. 105236.

COMPARATIVE EVALUATION OF METHODS FOR COMPENSATING METEOROLOGICAL EFFECTS ON IMAGES OF WILD ANIMALS

PhD stud. Natalenko D.N.

Reshetnev Siberian State University of Science and Technology, Institute of Informatics and Telecommunications

The paper discusses methods for compensating meteorological effects such as rain, snowfall, and fog on images of wild animals. Given the importance of accurate visual data for ecology and wildlife conservation, traditional image processing methods and modern neural network approaches are presented. The evaluation of their effectiveness is based on experimental data and image quality metrics, allowing for the identification of the strengths and weaknesses of each approach. The results of the study may be useful for specialists in computer vision, ecology, and wildlife conservation, as well as for software developers involved in image processing. The paper emphasizes the importance of selecting the optimal method for improving the analysis of images of wild animals.

ОЦЕНКА КАЧЕСТВА СЕГМЕНТАЦИИ В ЗАДАЧЕ ОБНАРУЖЕНИЯ МАСКИРУЮЩИХСЯ ОБЪЕКТОВ ЖИВОЙ ПРИРОДЫ

асп. Саидов Н.В.[1]

[1] Сибирский государственный университет науки и технологий имени академика М. Ф. Решетнева

Обнаружение маскирующихся объектов представляет собой сложную задачу в области компьютерного зрения, связанную с выявлением объектов, которые визуально сливаются с окружающей средой. В данной статье рассматривается проблема оценки качества сегментации маскирующихся объектов в живой природе. Приводится использование четырёх ключевых метрик: средней абсолютной ошибки (M), F-меры (F_β), E-меры ($E\phi$) и структурной меры ($S\alpha$), которые позволяют количественно оценить качество сегментации. Представлены формулы расчёта каждой метрики и проанализированы их преимущества в оценке замаскированных объектов. В экспериментальной части проводится сравнительный анализ с использованием различных наборов данных изображений дикой природы. Полученные результаты подчеркивают важность правильной оценки качества сегментации для повышения эффективности обнаружения маскирующихся объектов.

Введение. Обнаружение маскирующихся объектов (Camouflaged Object Detection, COD) является важной и сложной задачей в области компьютерного зрения. Маскирующиеся объекты стремятся сливаться с окружающей средой, что делает их трудноразличимыми даже для человеческого глаза. Задача автоматического обнаружения таких объектов представляет особый интерес как для научного сообщества, так и для прикладных областей, таких как биология, экология, медицина. Целью работы является рассмотрение проблемы оценки качества сегментации маскирующихся объектов в живой природе.

Основная проблема в решении задачи COD заключается в сложности сегментации, где объекты практически неотличимы от фона. В связи с этим становится необходимым использование эффективных метрик для оценки сегментации. В данной статье проводится исследование четырех ключевых метрик, которые позволяют количественно оценивать качество сегментации: средняя абсолютная ошибка (M), F-мера (F_β), E-мера ($E\phi$) и структурная мера ($S\alpha$). Эти метрики обеспечивают всестороннюю оценку точности и структурного сходства предсказанных масок с эталонными.

1. Оценки качества сегментации. Для начала рассмотрим теоретические аспекты метрик. Введём основные формулы для их расчёта, а также попытаемся объяснить принцип их работы. Перейдем к рассмотрению оценок качества сегментации (метрик).

Средняя абсолютная ошибка M является общеизвестной метрикой и измеряет среднюю разницу между предсказанными значениями маски и эталонными значениями. Эта метрика оценивает, насколько точно предсказанная маска совпадает с эталонной маской по всем пикселам изображения [1]. Формула средней абсолютной ошибки выглядит следующим образом:

$$M = \frac{1}{N}\sum_{i=1}^{N}|P_i - G_i|,$$

где N – общее количество пикселов в изображении; P_i – предсказанное значение интенсивности пиксела i в предсказанной маске; G_i – истинное значение интенсивности пиксела i в эталонной (ground truth) маске. Алгоритм работы средней абсолютной ошибки, в общем случае, сводится к следующему: сначала для каждого пиксела вычисляется абсолютная разница между предсказанным значением и эталонным, затем все разницы суммируются. Полученная сумма делится на общее количество пикселов N, чтобы получить среднюю ошибку. Если M равна нулю, то между эталонной и предсказанной маской нет абсолютно никаких различий (идеальное совпадение). Если M больше нуля, то существуют различия между масками (чем выше M, тем различия более существенные). Когда средняя абсолютная ошибка достигает единичного значения, то можно сделать вывод, что маски изображения абсолютно разные, что является самой худшей ситуацией для обнаружения замаскированного объекта.

Чтобы понять принцип работы данной метрики на практике рассмотрим два простых примера.

Пример 1. Имеются матрицы пикселов эталонной и предсказанной маски изображения.

$$P = [1, 0, 1, 0], \qquad G = [1, 0, 1, 0].$$

В данном случае средняя абсолютная ошибка M будет равна нулю, так как все пикселы в каждой позиции совпадают.

Пример 2. Также имеются матрицы пикселов изображения, однако значения пикселов в определённых позициях будут различаться.

$$P = [1, 0, 1, 0], \qquad G = [0, 1, 0, 1].$$

Рассчитаем среднюю абсолютную ошибку:

$$M = \frac{|1-0| + |0-1| + |1-0| + |0-1|}{4} = \frac{4}{4} = 1.$$

Как видно, средняя абсолютная ошибка оказалась равной единице, что указывает на абсолютное различие масок. Здесь стоит отметить, что на практике значение M равное нулю, как правило, невозможно, поскольку нет таких систем обнаружения, которые смогли бы идеально построить одинаковые эталонную и предсказанную маски.

К очевидным преимуществам средней абсолютной ошибки можно отнести простоту вычислений, поскольку формула основывается на простых арифметических операциях. Среднюю абсолютную ошибку легко логически интерпретировать по причине того, что она применима для любых масок. Однако данная метрика имеет ряд существенных недостатков, которые стали основанием для реализации других метрик. Одним из существенных недостатков является игнорирование пространственных зависимостей. Оценивается только средняя разница между пикселами, не принимая во внимание их пространственные отношения. Все ошибки имеют одинаковый вес. Это означает, что небольшие ошибки обрабатываются так же, как и крупные, что может быть нежелательно в задачах, где важны именно большие отклонения. Также данная метрика не учитывает форму объектов и качество их границ. Маска с хорошо выделенными границами и маска с размытыми краями могут иметь одинаковое значение средней абсолютной ошибки.

Перейдём к рассмотрению F-меры – метрики, объединяющей две важные характеристики качества предсказаний: точность (precision) и полноту (recall)**.** Она используется для оценки баланса между этими двумя аспектами, особенно в задачах бинарной классификации и сегментации [2]. Введём основную формулу для расчёта F-меры:

$$F_\beta = \frac{(1+\beta^2)\cdot \mathrm{P}\cdot \mathrm{R}}{\beta^2\cdot \mathrm{P} + \mathrm{R}},$$

где P (precision) – доля истинно положительных предсказаний среди всех положительных предсказаний; R (recall) – доля истинно положительных предсказаний среди всех реальных положительных значений; β – параметр, который определяет вес полноты по сравнению с точностью. Если β равен единице, то F-мера становится сбалансированной между точностью и полнотой.

Точность (precision) можно рассчитать по следующей формуле:

$$P = \frac{TP}{TP + FP},$$

где TP – число истинно положительных исходов; FP – число ложноположительных исходов.

Полнота (recall) рассчитывается по следующей формуле:

$$R = \frac{TP}{TP + FN},$$

где *TP* – число истинно положительных исходов; *FN* – число ложноотрицательных исходов.

Допустим, что в нашем сценарии число истинно положительных исходов (*TP*) равно 80, число ложноположительных исходов (*FP*) равно 20, число ложноотрицательных исходов (*FN*) – 10. Тогда F-мера рассчитывается, как:

$$F_\beta = \frac{(1 + 1^2) \cdot 0{,}8 \cdot 0{,}8889}{1^2 \cdot 0{,}8 + 0{,}8889} \approx 0{,}842.$$

F-мера даёт сбалансированное представление о производительности модели, учитывая как точность, так и полноту. Параметр β позволяет регулировать важность точности или полноты в зависимости от конкретной задачи. На основании полученных расчётов можно сказать, что F-мера хорошо подходит для задач, где объекты встречаются редко, так как она учитывает как ложные срабатывания, так и пропущенные объекты.

Теперь перейдём к рассмотрению *E*-меры – метрики, которая оценивает сходство между предсказанной и эталонной масками, учитывая как локальные, так и глобальные различия. Она предназначена для измерения структурного сходства, обеспечивая более глубокий анализ качества сегментации, чем простое совпадение пикселей [3]. Основная формула для расчёта выглядит следующим образом:

$$E_\phi = \frac{1}{N}\sum_{i=1}^{N} \phi(P_i\,,\, G_i),$$

где N – общее количество пикселов; P_i – предсказанное значение пиксела i; G_i – истинное значение пиксела i; $\phi(P_i, G_i)$ – функция, которая измеряет визуальное сходство между предсказанным значением P_i и эталонным значением G_i.

Функция ϕ направлена на оценку локального визуального сходства между пикселами предсказанной и эталонной масок. Она может учитывать различные аспекты, такие как разница в яркости, пространственное положение и локальный контекст.

При расчёте функции ϕ можно использовать следующую формулу:

$$\phi(P_i\,,\, G_i) = \frac{2 \cdot P_i \cdot G_i}{P_i^2 + G_i^2 + \epsilon},$$

где ϵ – небольшое положительное значение, предотвращающее деление на ноль.

Далее покажем наглядно принцип работы этой метрики. Пусть имеется матрица P, которая содержит значения предсказанной маски, а также матрица G, содержащая значения эталонной маски:

$$P = [1, 0, 1, 0], \qquad G = [0{,}9, 0{,}1, 0{,}8, 0{,}2].$$

Рассчитаем *E*-меру, используя функцию ϕ. Для расчёта берём каждое значение из каждой матрицы соответственно. Для примера рассчитаем ϕ_1 и ϕ_2. Значения ϕ_3 и ϕ_4 аналогично:

$$\phi_1(1\,,\, 0{,}9) = \frac{2 \cdot 1 \cdot 0{,}9}{1^2 + 0{,}9^2 + \epsilon} \approx \frac{1{,}8}{1{,}81} \approx 0{,}9945.$$

$$\phi_2(0\,,0{,}1) = \frac{2 \cdot 0 \cdot 0{,}1}{0^2 + 0{,}1^2 + \epsilon} = 0.$$

Получаем итоговое значение *E*-меры:

$$E_\phi = \frac{0{,}9945 + 0 + 0{,}9756 + 0}{4} = \frac{1{,}9701}{4} \approx 0{,}4925.$$

E-мера учитывает не только точность совпадения пикселов, но и общее структурное сходство масок, включая их форму и расположение. Можно утверждать, что в задачах обнаружения маскирующихся объектов, где границы объектов могут быть размыты или трудно различимы, *E*-мера помогает оценить качество сегментации более объективно, хотя сильно зависит от качества эталонной маски.

Рассмотрим последнюю метрику – структурная мера (S_α), которая оценивает сходство между предсказанной и эталонной масками с учётом как внутренней структуры объекта, так и окружающего фона. Она специально разработана для измерения структурного согласия между предсказаниями и эталонными масками. Введём основную формулу для расчёта этой метрики:

$$S_\alpha = \alpha \cdot S_{object} + (1 - \alpha) \cdot S_{region},$$

где S_{object} – структурная мера объекта; S_{region} – структурная мера для региона (фон); α – весовой коэффициент (обычно α = 0.5), который балансирует вклад между объектом и регионом.

Структурная мера объекта оценивает, насколько предсказанная маска совпадает с эталонной в области объекта. Структурную меру объекта можно рассчитать по следующей формуле:

$$S_{object} = \frac{1}{N_0} \sum_{i \epsilon O} \phi(P_i\,,\, G_i),$$

где N_0 – количество пикселов объекта; $\phi(P_i\,,\, G_i)$ – функция оценки локального сходства (как в *E*-мере); *O* – область объекта в эталонной маске.

Региональную меру можно рассчитать подобным образом:

$$S_{region} = \omega \cdot S_{fg} + (1 - \omega) \cdot S_{bg},$$

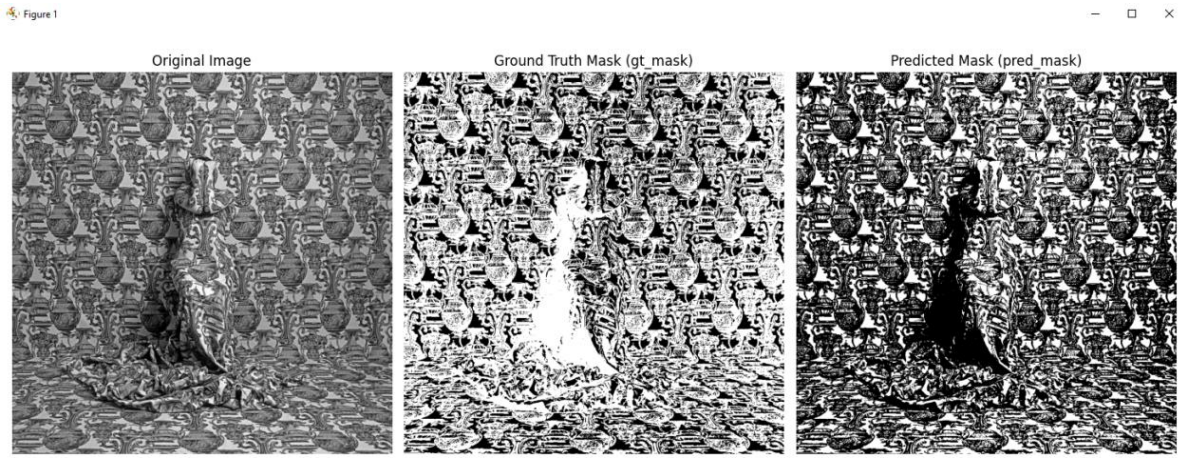
где ω – доля пикселей объекта в эталонной маске; S_{fg} – мера для переднего плана (объекта); S_{bg} – мера для заднего плана (фона).

Структурная мера оценивает, насколько предсказанная маска корректно охватывает объект, учитывая локальное сходство между пикселами объекта [4]. Если S_α равна единице, то мы получаем идеальное совпадение предсказанной и эталонной масок. Если S_α равно или близко к нулю, то маски разные, что указывает на плохую сегментацию изображения.

К преимуществам структурной меры можно отнести возможность учёта как структуры фона, так и структуры объекта конкретного изображения. Регулируя параметр α можно выбирать что для изображения будет более важно – фон или объект. Однако, по сравнению с другими метриками, показатель S_α требует более сложных вычислений.

Подводя итог, можно отметить, что S_α позволяет оценивать качество сегментации с учётом структурных особенностей объектов и фона. Она особенно полезна в задачах, где важна не только точность предсказания, но и качество выделения объектов и их границ [5]. Использование S_α в сочетании с другими метриками обеспечивает более полное понимание эффективности моделей сегментации.

2. Экспериментальная часть. В рамках исследования была разработана программа на языке Python, которая подсчитывает описанные метрики. На рисунке 1 демонстрируется формирование программой эталонной и предсказательной масок для исходного изображения.



а) б) в)

Рисунок 1 – Формирование эталонной и предсказательной масок (а – оригинальное изображение, б – эталонная маска, в – предсказательная маска)

Поскольку загруженное в программу изображение не подвергалось методам машинного обучения, то можно увидеть, насколько маски изображения отличаются. После запуска программы на расчёт оценок были получены следующие результаты (таблица 1).

Таблица 1. Полученные оценки исходного изображения

Метрика	**Результат**
Средняя абсолютная ошибка M	1,000
F-мера	0,000
E-мера	0,186
Структурная мера S_α	0,395

Далее представлены результаты масок изображений с набора данных CHAMELEON [6], полученных с применением нейронных сетей SpectralSeg, BigGW, TokenCut (рисунок 3).

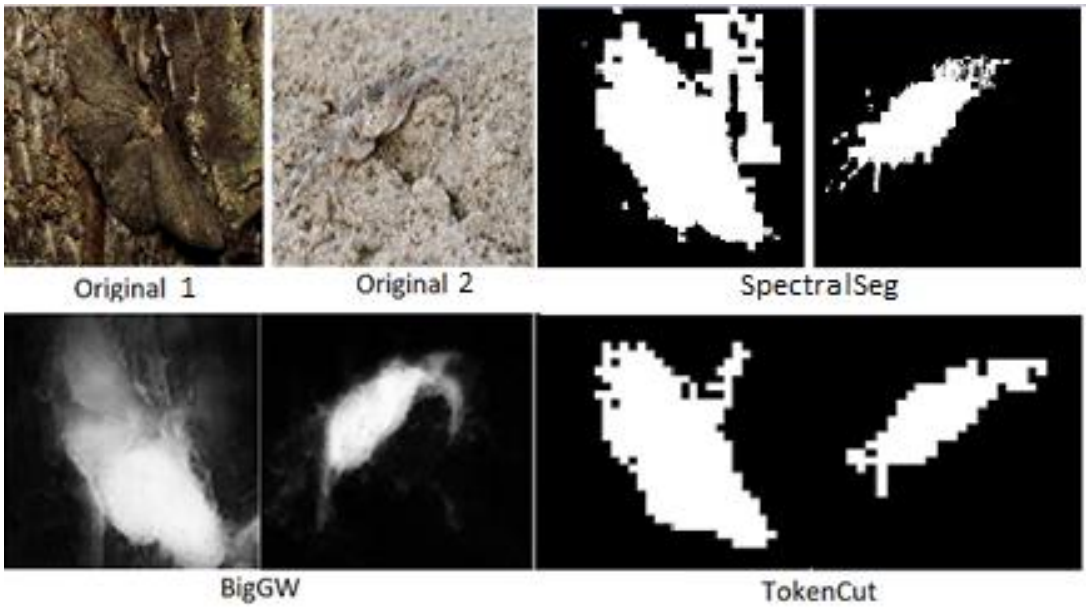


Рисунок 3 – Маски изображений, полученные с использованием нейронных сетей

В таблице 2 представлены результаты полученных метрик для данных изображений. Как можно увидеть, точность совпадения масок значительно выше. Стоит отметить, что в качестве оценок взяты средние значения двух исходных изображений.

Таблица 2. Рассчитанные результаты метрик, полученных масок изображений

Метрика	SpectralSeg	BigGW	TokenCut
Средняя абсолютная ошибка M	0,220	0,257	0,132
F-мера	0,440	0,294	0,536
E-мера	0,628	0,527	0,740
Структурная мера S_α	0,575	0,547	0,654

Заключение. Рассмотренные в работе метрики играют ключевую роль в оценке качества сегментации замаскированных объектов. Каждая из этих метрик обладает уникальными характеристиками и предоставляет разные перспективы для анализа эффективности моделей обнаружения. Средняя абсолютная ошибка позволяет оценить среднее отклонение предсказанных масок от эталонных, что обеспечивает простую и интуитивно понятную оценку точности предсказаний. Однако её ограниченность в учёте пространственных структур требует использования дополнительных метрик. Что касается F-меры, то она особо значима в задачах с несбалансированными данными, где ошибки пропуска объектов и ложных срабатываний имеют разную значимость. E-мера, за счёт использования функции ϕ, обеспечивает детальную оценку локального и глобального сходства между предсказанными и эталонными масками. Структурная мера объединяет оценку объекта и фона, обеспечивая более комплексный подход к оценке качества сегментации. В совокупности, использование всех этих метрик позволяет проводить всестороннюю оценку моделей обнаружения замаскированных объектов, обеспечивая глубокий анализ их сильных и слабых сторон.

Литература

1. He C. et al. Strategic preys make acute predators: Enhancing camouflaged object detectors by generating camouflaged objects //arXiv preprint arXiv:2308.03166. – 2023.
2. M. Zhuge. et al. Salient object detection via integrity learning. IEEE Trans. Pattern Anal. Mach. Intell., 2022.
3. X. Hu. et al. High-resolution Iterative Feedback Network for Camouflaged Object Detection // arXiv preprint arXiv:2203.11624. 2022.
4. L. Deng. et al. A noise robust conditional generative adversarial network for one shot learning. IEEE Trans. In tell. Transp. Syst., 23(12): 25249–25258, 2022.
5. Офицеров В. А., Конушин А. С. Нейросетевые методы сегментации изображений высокого разрешения //International Journal of Open Information Technologies. – 2024. – Т. 12. – №. 6. – С. 57-64.
6. Chameleon Dataset, Deep Graph Library. January 03, 2025, [Online]. Available: https://docs.dgl.ai/en/2.0.x

ASSESSMENT OF THE QUALITY OF SEGMENTATION IN THE TASK OF DETECTING CAMOUFLAGED WILDLIFE OBJECTS

PhD stud. Saidov N.V.[1]

[1] Reshetnev Siberian State University of Science and Technology

Camouflaged Object Detection (COD) is a complex task in the field of computer vision related to the identification of objects that visually blend into the environment. This paper discusses the problem of estimating the segmentation quality in the context of detecting such objects in wildlife. Four key metrics are used: mean absolute error (M), F-measure (F_β), E-measure ($E\phi$) and structural measure (S_α), which allow quantifying the quality of segmentation. Formulas for calculating each metric are presented and their advantages in evaluating disguised objects are analyzed. In the experimental part, a comparative analysis is carried out using different datasets of wildlife images. The results obtained emphasize the importance of correctly estimating the quality of segmentation in order to increase the effectiveness of camouflaged object detection.

СВЕРТОЧНЫЕ НЕЙРОННЫЕ СЕТИ В ЗАДАЧЕ АВТОМАТИЧЕСКОГО РАСПОЗНАВАНИЯ ЦИФРОВЫХ ВИДОВ МОДУЛЯЦИИ

асп. Савин С.А., асп. Усков Н.А., доц. Красавин К.С., студ. Соколова Д.Д.

Ярославский государственный университет им. П.Г. Демидова

Разработана модель сверточной нейронной сети (СНС), позволяющая добиться средней точности распознавания 95-99% при распознавании цифровых модуляций BPSK, QPSK и 16QAM. Для обучения и тестирования нейросетевого алгоритма используются данные из базы модулированных сигналов DeepSig. Входными данными системы являются отсчеты I/Q компонент радиосигнала. На ее выходе распознается вид цифровой модуляции входного сигнала со средней точностью 98,44% при отношении сигнал/шум от 0 дБ до 20 дБ. Представлены графики потерь и точности, которые показывают зависимость данных характеристик от количества эпох обучения сверточной нейронной сети. Также получена матрица ошибок, где можно увидеть и проанализировать ошибки рассмотренной нейросетевой модели при распознавании указанных видов цифровой модуляции радиосигналов.

Введение

Современные радиотехнические системы широко применяются в различных областях, включая связь, радиолокацию и беспроводные сети передачи данных. Одним из важнейших аспектов таких систем является выбор типа модуляции, который влияет на эффективность передачи и обработки информации. Автоматическое распознавание видов модуляции играет критически важную роль в обеспечении надежности связи, оптимизации спектрального ресурса и обеспечении безопасности радиокоммуникаций [1-2].

Традиционные методы распознавания модуляции основаны на использовании классических алгоритмов обработки сигналов в данной области, таких как метод моментов, анализ статистических характеристик и методы на основе эвристик. Однако данные подходы обладают рядом ограничений, в частности, они достаточно чувствительны к изменению характеристик канала связи, шумам и помехам. В связи с этим, возникает необходимость в использовании более устойчивых методов, таких как сверточные нейронные сети, которые могут автоматически выделять значимые признаки сигналов и адаптироваться к сложным помеховым условиям приема сигналов [3-4].

Сверточные нейронные сети представляют собой мощный инструмент, позволяющий значительно улучшить точность и скорость распознавания видов модуляции, в том числе и цифровой. Они способны выделять локальные иерархические признаки сигналов, что делает их эффективными для задач их классификации [5].

Методы и алгоритмы

Для обучения СНС использован широко распространенный датасет RadioML2018.01A, содержащий различные типы аналоговых и цифровых модуляций, в том числе широко используемые во многих стандартах современных беспроводных систем связи BPSK, QPSK и 16QAM. Структура набора данных RadioML 2018.01A показана на рисунке 1. Данный набор содержит 2 555 904 кадра, каждый из которых представляет собой 1024 выборки сложных временных рядов. Эти выборки состоят из синфазных (I) и квадратурных (Q) составляющих, представленных в виде чисел с плавающей запятой.

Всего в базе представлены 24 вида аналоговой и цифровой модуляции, включая как классические, так и более современные: OOK, 4ASK, 8ASK, BPSK, QPSK, 8PSK, 16PSK, 32PSK, 16APSK, 32APSK, 64APSK, 128APSK, 16QAM, 32QAM, 64QAM, 128QAM, QAM256, AM_SSB_WC, AM_SSB_SC, AM_DSB_WC, AM_DSB_ SC, FM, GMSK, OQPS. В данном исследовании рассматривались только цифровые виды модуляции, которые, как уже отмечалось ранее, достаточно широко применяются в современных беспроводных системах связи: BPSK, QPSK и 16QAM.

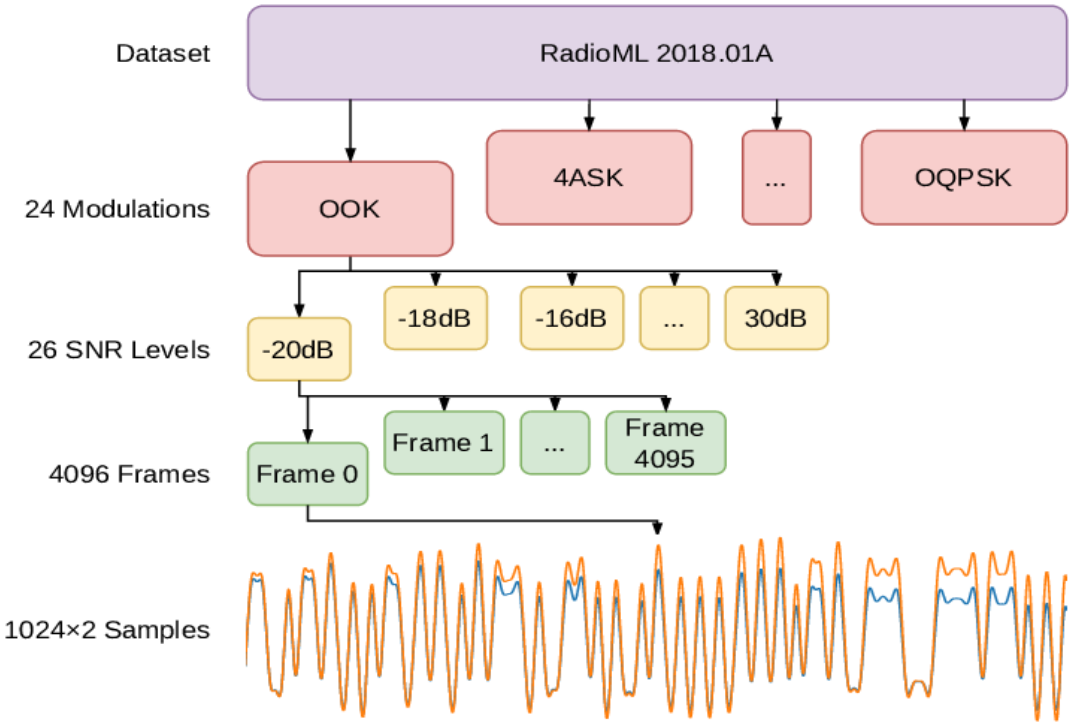


Рисунок 1 – Структура набора данных RadioML 2018.01A

Обучение сетей

В работе использована модифицированная сверточная нейронная сеть CNN_NET. Архитектура данной сети включает в себя входной слой, сверточный слой, слой пулинга, полносвязные слои. Кроме того, в данной сети применены специализированные сверточные блоки PGIBlock и GELANBlock. Данные блоки обеспечивают захват более сложных зависимостей в данных благодаря использованию нескольких сверточных слоев и их конкатенации. На рисунке 2 показана структура модифицированной сверточной нейронной сети CNN_NET.

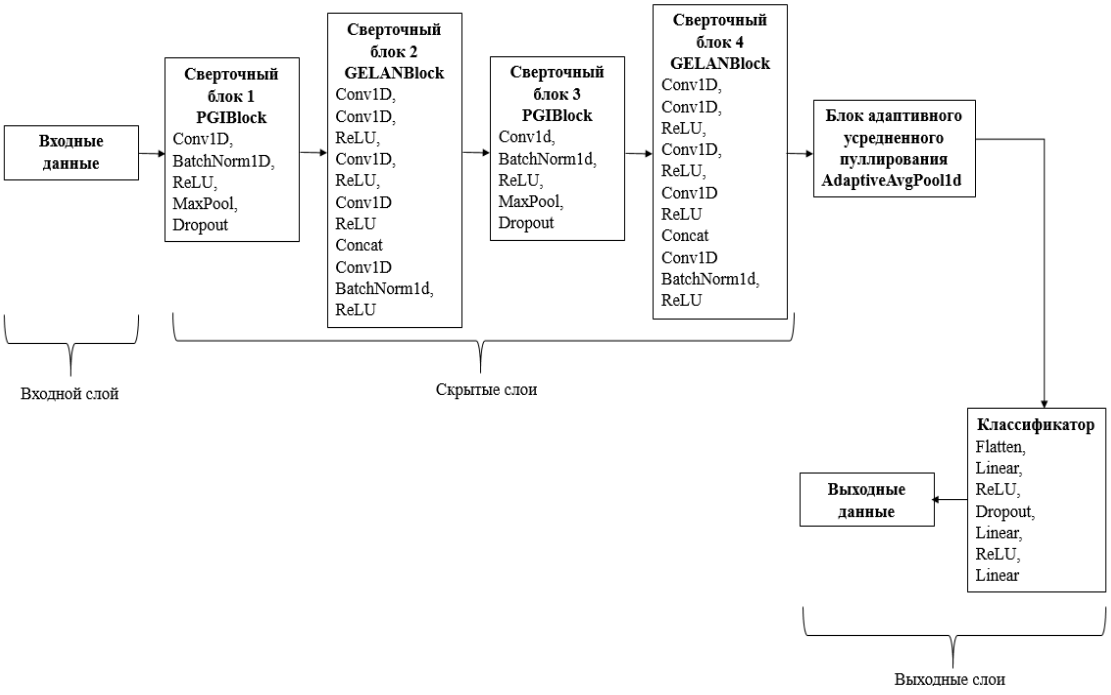


Рисунок 2 – Структура модифицированной нейронной сети CNN_NET

Обучение СНС проводилось с использованием батчей размером 64, а количество эпох составило 100.

В компьютерном эксперименте отношение сигнал/шум (E_s/N_0) варьировалось от 0 дБ до 20 дБ с шагом 2 дБ, что во многом соответствует реальным применениям указанных видов цифровой модуляции в современных беспроводных системах и сетях связи. Для оптимизации параметров нейросетевой модели использован алгоритм Adam с начальным коэффициентом обучения 0,001.

Анализ полученных результатов

Установлено, что процесс обучения сверточной нейронной сети характеризовался монотонным уменьшением значения функции потерь, что свидетельствует об успешной сходимости алгоритма оптимизации. График соответствующих потерь представлен на рисунке 3. Зависимость точности распознавания от количества эпох на рисунке 4 показывает, что увеличение точности достигает 98,44%.

Полученные метрики подтверждают высокую обобщающую способность разработанной архитектуры и её устойчивость к переобучению. Анализ кривых обучения и валидации не выявил признаков деградации модели, что подчеркивает корректность выбора гиперпараметров и процедуры регуляризации.

Матрица ошибок, представленная на рисунке 5, показывает, что модель точно распознает большинство видов цифровой модуляции. Однако наибольшие ошибки наблюдаются между модуляциями 16QAM и QPSK, что указывает на их определенную похожесть друг на друга и, соответственно, на сложности в их различении. Более конкретно – эти ошибки могут быть связаны с высокой степенью схожести их спектральных характеристик, а также с возможным наличием дополнительных шумовых факторов в обучающих данных.

Установлено, что в целом, рассматриваемая нейросетевая система демонстрирует достаточно высокую надежность и устойчивость к шуму в указанных пределах изменения ОСШ.

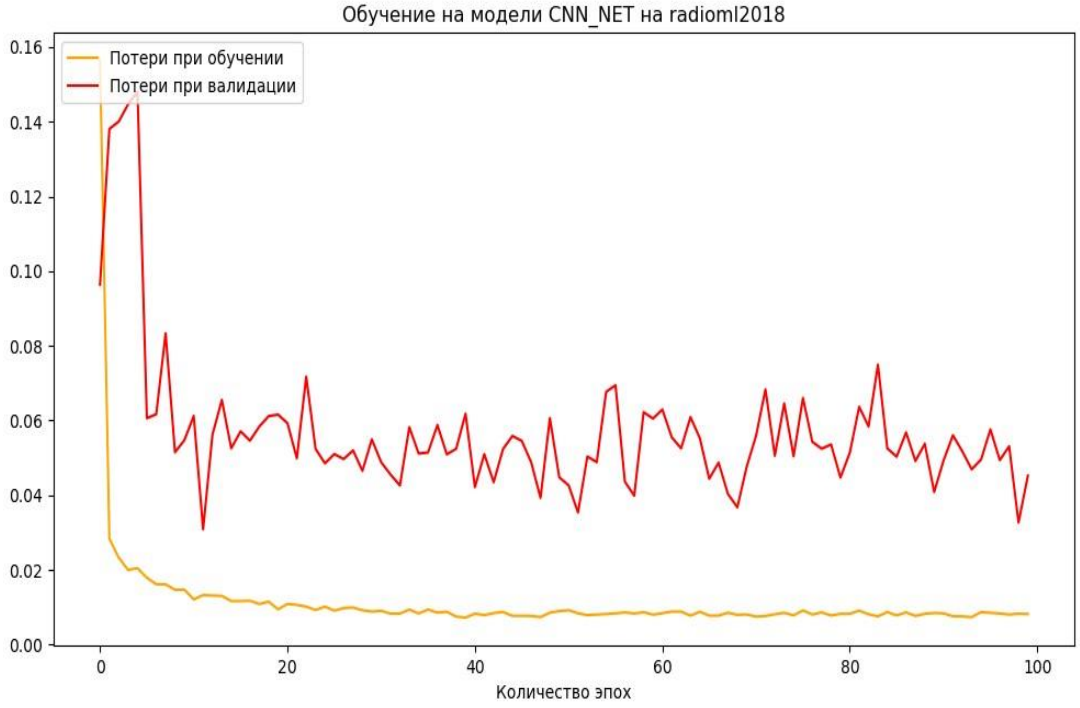


Рисунок 3 – График потерь сверточной нейронной сети CNN_NET

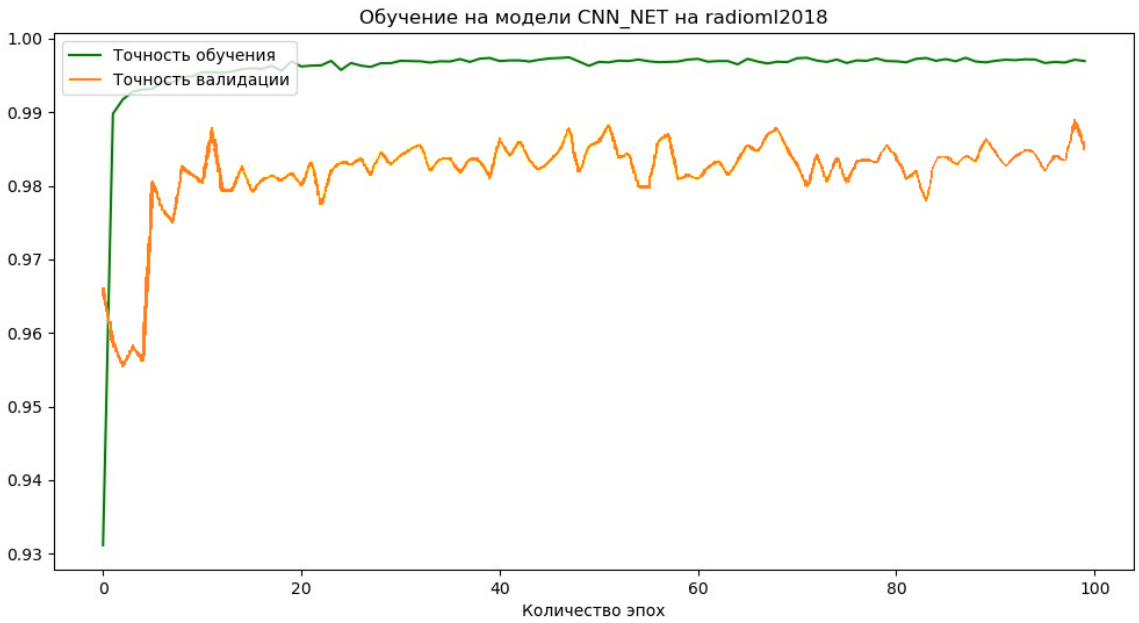


Рисунок 4 – График точности сверточной нейронной сети CNN_NET

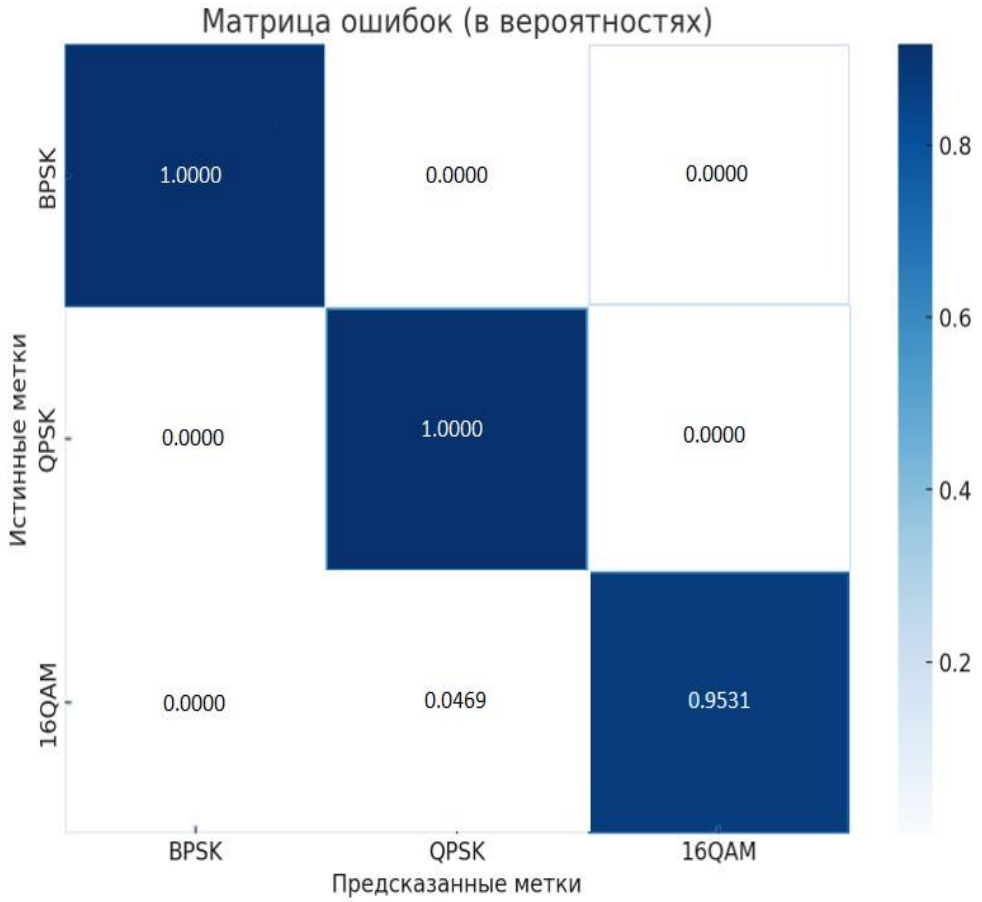


Рисунок 5 – Матрица ошибок сверточной нейронной сети CNN_NET

Заключение

Применение сверточной нейронной сети указанного вида в задаче автоматического распознавания цифровых видов модуляции показало свою эффективность на примере следующих цифровых модуляций: BPSK, QPSK и 16QAM, широко использующихся в настоящее время в различных стандартах беспроводных сетей связи различного назначения. Входными данными данной СНС являются отсчеты I/Q компонент соответствующих радиосигналов.

Для обучения СНС использована популярная база радиосигналов RadioML2018.01A, содержащая различные типы цифровых и аналоговых модуляций, в том числе и исследуемые в данной работе.

Разработанная архитектура нейронной сети CNN_NET с модифицированными сверточными блоками PGIBlock и GELANBlock позволила добиться средней точности распознавания 98,44% в диапазоне отношений сигнал/шум от 0 до 20 дБ с шагом 2 дБ.

Литература

1. Парамонов, А.А. Многозадачная нейронная сеть в задаче распознавания вида qam- и psk-модуляции в условиях параметрической априорной неопределенности / Парамонов А.А., Нгуен В.М., Нгуен М.Т. // Russian Technological Journal. 2023. Т. 11. № 4. С. 49-58.
2. Nandi, A. Automatic Modulation Classification: Principles, Algorithms and Applications. Wiley / A. Nandi, Z. Zhu. 2018. 194 с.
3. Беляева, М.Б. Распознавание модуляции: что можно узнать, если заранее ничего не известно // Цифровая обработка сигналов. 2013. № 2. С. 55-63.
4. Аджемов, С.С. Методы распознавания видов цифровой модуляции сигналов в когнитивных радиосистемах / С.С. Аджемов, Н.В. Кленов, М.В. Терешонок, Д.С. Чиров // Вестник Московского университета. Серия 3. Физика. Астрономия. 2015. № 6. С. 19-27.
5. Крон, Дж. Глубокое обучение в картинках. Визуальный гид по искусственному интеллекту / Дж. Крон, Г. Бейлевельд, Б. Аглаэ. СПб.: Питер, 2020. 400 с.

CONVOLUTIONAL NEURAL NETWORKS IN THE TASK OF AUTOMATIC RECOGNITION OF DIGITAL TYPES OF MODULATION

Savin S.A., Uskov N.A., Krasavin K.S., Sokolova D.D.

P.G. Demidov Yaroslavl State University

A convolutional neural network (SNN) model has been developed that allows achieving an average recognition accuracy of 95-99% when recognizing digital modulations BPSK, QPSK and 16QAM. Data from the Deepsign modulated signal database is used to train and test the neural network algorithm. The input data of the system are samples of the I/Q components of the radio signal. Its output detects a type of digital modulation of the input signal with an average accuracy of 98.44% with a signal-to-noise ratio from 0 dB to 20 dB. Loss and accuracy graphs are presented, which show the dependence of these characteristics on the number of training epochs of a convolutional neural network. An error matrix has also been obtained, where it is possible to see and analyze the errors of the considered neural network model when recognizing these types of digital modulation of radio signals.

ВИЗУАЛЬНО-ТЕКСТОВЫЙ ДЕСКРИПТОР ДЛЯ ПОВТОРНОЙ ИДЕНТИФИКАЦИИ ЧЕЛОВЕКА НА ИЗОБРАЖЕНИИ

асп. Савков М.В.

Федеральное государственное бюджетное образовательное учреждение высшего образования «Сибирский государственный университет науки и технологий имени академика М.Ф. Решетнева»

Повторная идентификация человека (Re-Identification, ReID) представляет собой одну из важнейших задач компьютерного зрения, особенно в контексте видеонаблюдения и анализа поведения. Изменения ракурса, освещения и внешнего вида объекта приводят к потере дискриминативных признаков, что снижает эффективность традиционных методов. В данной статье представлен мультимодальный подход, основанный на визуально-текстовом дескрипторе, который объединяет визуальные данные с текстовыми описаниями объектов. Экспериментальные результаты показывают, что использование предложенного подхода улучшает точность идентификации по сравнению с современными методами.

Введение

Системы повторной идентификации человека находят применение в различных областях автоматизации систем видеонаблюдения и анализа поведения. ReID представляет собой сложную задачу, связанную с установлением соответствия между изображениями одного и того же человека, снятыми в различных условиях. Основные сложности включают изменения ракурса, освещения, внешнего вида и окружения [1, 2].

Традиционные подходы к ReID основываются на использовании глубоких сверточных нейронных сетей (CNN), таких как ResNet [3], которые способны извлекать дискриминативные признаки из изображений. Эти методы показывают высокую точность на статичных наборах данных, таких как Market-1501 и DukeMTMC-reID, однако их эффективность снижается в реальных условиях, где присутствуют значительные изменения ракурса или освещения [4].

Атрибутивные методы, использующие текстовые описания, такие как цвет одежды или наличие аксессуаров, предлагают дополнительный уровень устойчивости. Например, работы [5, 6] продемонстрировали, что текстовые атрибуты могут компенсировать недостатки визуальных признаков, особенно в условиях сильных изменений угла обзора или частичной обструкции. Тем не менее, такие подходы требуют ручной аннотации данных, что ограничивает их масштабируемость.

Мультимодальные методы, такие как CLIP [5], объединили визуальные признаки с текстовыми описаниями, используя контрастивное обучение. Эти методы обеспечивают высокий уровень переносимости моделей между различными задачами и демонстрируют устойчивость к изменениям условий. Однако их применение к задаче ReID остается недостаточно изученным.

Целью данной работы является разработка и исследование мультимодального подхода, основанного на визуально-текстовом дескрипторе (ВТД), который объединяет визуальные и текстовые признаки для повышения точности ReID. Основной акцент делается на использовании текстовых данных для компенсации потерь информации, вызванных изменением ракурса.

Визуально-текстовый дескриптор

Для разработки и оценки мультимодальных моделей, применяемых к задаче повторной идентификации человека, используются разнообразные наборы данных, которые обеспечивают моделям возможность справляться с изменчивыми условиями съемки. Эти наборы данных включают как изображения из реальных условий, так и сложные сценарии с вариациями ракурсов, освещения и обстановки. Наиболее значимыми наборами данных в этой области являются:

– Market-1501: содержит 32 668 изображений 1 501 человека, снятых с шести камер. Разнообразие ракурсов обеспечивает реалистичное тестирование моделей [2];

– DukeMTMC-reID: включает 36 411 изображений 1 404 человека, дополненных "ложными" изображениями (distractors), которые создают условия, близкие к реальным задачам ReID [2];

– CUHK03: насчитывает 13 164 изображения 1 467 человек, предлагая два типа аннотаций – ручную и автоматическую. Это позволяет тестировать модели на различных уровнях сложности [3];

– MSMT17: включает более 126 441 изображений 4 101 человека, снятых с 15 камер. Он предлагает разнообразие сцен и условий, включая вариации освещения и ракурсов, делая его наиболее обширным и сложным датасетом для ReID [5].

Примеры изображений из наборов данных показаны на рисунке 1.



а) б) в) г)

Рисунок 1 – Пример изображений наборов данных: а) Market-1501, б) DukeMTMC-reID, в) CUHK03, г) MSMT17

Эти наборы данных обеспечивают надежную основу для разработки визуально-текстового дескриптора, позволяя моделям работать в условиях, близких к реальным.

Методология разработки визуально-текстового дескриптора основана на интеграции двух основных компонентов – визуального и текстового. Визуальный компонент использует модифицированную архитектуру ResNet-50, которая позволяет извлекать дискриминативные признаки из изображений. Для повышения устойчивости к изменениям ракурса и обстановки модель обучается с использованием триплетной функции потерь (Triplet Loss). Текстовый дескриптор является важной составляющей предложенного подхода, так как он позволяет эффективно использовать семантическую информацию, извлечённую из текстовых описаний, для повышения точности повторной идентификации человека (ReID). Основная идея заключается в дополнении визуальных признаков, которые могут быть утеряны в сложных условиях съёмки, с помощью текстовых данных. Для реализации текстового дескриптора в данной работе используется модель BERT (Bidirectional Encoder Representations from Transformers) [6], которая показала высокую эффективность при обработке текстовых данных.

Процесс работы текстового дескриптора начинается с предварительной обработки текста. На этом этапе все текстовые описания объектов, такие как «чёрная куртка» или «синий рюкзак», нормализуются: удаляются лишние символы, текст приводится к нижнему регистру и исключаются стоп-слова. После этого текст токенизируется, преобразуясь в последовательность векторов, которые затем обрабатываются моделью BERT. Этот этап позволяет извлечь контекстуально-зависимые семантические представления, где каждый вектор учитывает как значение слова, так и его окружение. Например, такие оттенки, как «светло-синий» и «тёмно-синий», будут различимы, что играет ключевую роль в задаче идентификации.

Текстовый дескриптор обладает рядом преимуществ, которые делают его незаменимым в задачах ReID. Во-первых, он обеспечивает инвариантность к изменениям ракурса и освещения, так как семантические признаки не зависят от условий съёмки. Во-вторых, текстовые данные помогают восполнить пробелы, возникающие при отсутствии ключевых визуальных признаков, например, при частичном перекрытии объекта. Кроме того, текстовый дескриптор демонстрирует высокую гибкость, так как может быть адаптирован для различных сценариев,

включая мультиязыковые описания или автоматическую генерацию текста на основе изображений.

Для объединения визуальных и текстовых данных используется механизм внимания (Attention Mechanism), который помогает выделять наиболее важные признаки для задачи идентификации. Комбинированная функция потерь, включающая кросс-энтропийную потерю (Cross-Entropy Loss) и контрастивную потерю (Contrastive Loss), обеспечивает эффективное обучение модели, помогая интегрировать данные обеих модальностей. Итоговый дескриптор представляет собой объединение визуальных и текстовых признаков, что делает его устойчивым к внешним изменениям [7-9].

Оценка эффективности предложенного дескриптора проводится на основе метрик Mean Average Precision (mAP) и Rank-1 Accuracy [10-14], которые широко используются для измерения точности моделей ReID. Сравнение результатов с базовыми подходами, такими как ResNet-50, и мультимодальными методами, включая CLIP, демонстрирует преимущество предлагаемого подхода за счёт использования текстовых данных. Это подтверждает, что интеграция текстовой информации помогает повысить точность и устойчивость модели в реальных условиях [15-16].

Экспериментальные данные

Для демонстрации эффективности предложенного визуально-текстового дескриптора (ВТД) был проведён сравнительный анализ с рядом современных методов. Оценка проводилась на наиболее распространённых наборах данных для задач ReID: Market-1501, DukeMTMC-reID, MSMT17 и CUHK03.

Сравнение было проведено по ключевым метрикам, таким как Mean Average Precision (mAP), которая измеряет точность совпадений для всех возможных пар запросов, и Rank-1 Accuracy, отражающей долю случаев, когда корректное совпадение занимает первое место в ранжированном списке. Результаты представлены в таблице 1.

Таблица 1. Сравнение ключевых метрик для выбранных методов

Метод	Market-1501 (mAP)	DukeMTMC-reID (mAP)	MSMT17 (mAP)	CUHK03 (mAP)	Rank-1 Accuracy (Market-1501)
ResNet-50	78,5%	67,3%	50,2%	65,3%	84,3%
CLIP (ViT-B/32)	85,6%	74,2%	56,8%	70,4%	88,5%
TF-CLIP [3]	87,1%	75,4%	58,2%	72,1%	90,2%
ВТД	88,2%	76,1%	60,5%	73,8%	91,5%

Как видно из таблицы, предложенный ВТД демонстрирует стабильное улучшение точности по сравнению с другими методами. На наборе данных CUHK03, где точность других методов варьируется от 65,3% до 72,1%, ВТД достигает 73,8%, показывая заметное преимущество благодаря интеграции текстовых и визуальных данных. Аналогичные результаты наблюдаются на других наборах данных: на Market-1501 mAP ВТД составляет 88,2%, что превосходит ResNet-50 на 9,7% и CLIP на 2,6%. На DukeMTMC-reID и MSMT17 ВТД также демонстрирует лучшие результаты, сохраняя разрыв на уровне 5–10%.

Особенно примечательным является то, как ВТД справляется в условиях, где традиционные методы сталкиваются с трудностями. Например, на наборе данных CUHK03, содержащем изображения с частичным перекрытием объектов и вариациями углов съёмки, ВТД показал высокую устойчивость. Это связано с тем, что текстовые описания, такие как «чёрный рюкзак» или «синий пиджак», позволяли модели компенсировать потерю визуальных признаков, вызванную отсутствием полного изображения объекта.

На наборе данных MSMT17, включающем изображения, снятые в условиях сильного освещения и теней, визуально-текстовый дескриптор также превосходил другие методы. Например, в сценариях, где визуальные признаки искажались из-за яркого контраста, текстовые данные, такие как «белая футболка с логотипом», обеспечивали дополнительную информацию,

необходимую для точной идентификации. В таких случаях традиционные методы, такие как ResNet-50, часто ошибались из-за отсутствия устойчивых визуальных паттернов.

Таким образом, визуально-текстовый дескриптор показывает себя как особенно полезный инструмент в условиях, где визуальные данные ограничены или искажены. Интеграция текстовых данных позволяет существенно повысить точность идентификации и справляться с реальными сложными условиями, в которых традиционные методы теряют свою эффективность.

Заключение

В данной работе рассмотрен подход к задаче повторной идентификации человека, основанный на использовании ВТД. Предложенный метод эффективно сочетает визуальные признаки с текстовыми данными, что позволяет компенсировать потери информации, вызванные изменением ракурсов, освещения и других внешних факторов. Экспериментальные результаты, полученные на широко используемых наборах данных, таких как Market-1501, DukeMTMC-reID, MSMT17 и CUHK03, подтверждают, что ВТД превосходит традиционные методы, включая ResNet-50 и мультимодальные подходы, такие как CLIP.

Особенность ВТД заключается в его способности интегрировать данные из различных модальностей с использованием механизмов внимания и комбинированных функций потерь. Это позволяет улучшить показатели точности, такие как Mean Average Precision (mAP) и Rank-1 Accuracy, на всех тестируемых наборах данных. Анализ показал, что ВТД демонстрирует стабильное улучшение производительности на 5–10% по сравнению с традиционными методами, сохраняя высокую устойчивость к изменениям условий съёмки.

Тем не менее, несмотря на преимущества, предложенный метод имеет определённые ограничения. Во-первых, эффективность работы ВТД сильно зависит от качества текстовых описаний. При недостаточной точности или неполных текстах модель может терять в точности. Кроме того, существует проблема с масштабируемостью системы, так как создание текстовых аннотаций для больших объёмов данных требует значительных усилий. Вдобавок, использование текстовых данных может увеличивать вычислительные затраты, что важно учитывать при применении модели в реальных приложениях, таких как системы видеонаблюдения в реальном времени. Также, модель может сталкиваться с трудностями при обработке сильно искажённых или абстрактных текстовых описаний, что ограничивает её универсальность.

Полученные результаты подчёркивают потенциал мультимодальных подходов в задаче ReID. В будущем планируется оптимизировать архитектуру ВТД для снижения вычислительных затрат, а также исследовать автоматическую генерацию текстовых данных для уменьшения зависимости от ручной аннотации. Это позволит ещё шире использовать предложенный метод в реальных условиях, таких как системы видеонаблюдения и анализ поведения.

Литература

1. K. He, X. Zhang, S. Ren, J. Sun. Deep residual learning for image recognition // CVPR. 2016. pp. 770-778.
2. L. Zheng, Y. Yang, Q. Tian. Scalable person re-identification: A benchmark // ICCV. 2015. pp. 1116-1124.
3. Y. Sun, L. Zheng, W. Deng, S. Wang. SVDNet for pedestrian retrieval // ICCV. 2018. pp. 3800-3808.
4. С.А. Игнатьева, Р.П. Богуш. Состав и особенности нового набора данных для решения задачи повторной идентификации человека в системах видеонаблюдения // Информатика: проблемы, методы, технологии: сборник материалов XXII международной научно-практической конференции. Воронеж. 2022. С. 592-599.
5. A. Radford, J. Kim, C. Hallacy, A. Ramesh. Learning transferable visual models from natural language supervision // arXiv:2103.00020. 2021.

6. Y. Wu, Y. Lin, X. Dong, Y. Yan, W. Ouyang. Exploit the Unknown Gradually: One-Shot Video-Based Person Re-Identification by Stepwise Learning // CVPR. 2019. pp. 5177-5186.
7. J. Devlin, M. Chang, K. Lee, K. Toutanova. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding // NAACL. 2019. pp. 4171-4186.
8. С.А. Игнатьева, Р.П. Богуш. Алгоритм реидентификации людей по изображениям систем видеонаблюдения с использованием нейросетевого составного дескриптора // Информатика: проблемы, методы, технологии: сборник материалов XXIII международной научно-практической конференции. Воронеж. 2023. С. 1077–1084.
9. S. Yan, N. Dong, L. Zhang, J. Tang. CLIP-Driven Fine-grained Text-Image Person Re-identification // arXiv:2210.10276. 2022.
10. C. Yu, X. Liu, Y. Wang, P. Zhang, H. Lu. TF-CLIP: Learning Text-free CLIP for Video-based Person Re-Identification // arXiv:2312.09627. 2023.
11. X. Yu, N. Dong, L. Zhu, H. Peng, D. Tao. CLIP-Driven Semantic Discovery Network for Visible-Infrared Person Re-Identification // arXiv:2401.05806. 2024.
12. H. Zhao, L. Qi, X. Geng. CLIP-DFGS: A Hard Sample Mining Method for CLIP in Generalizable Person Re-Identification // arXiv:2410.11255. 2024.
13. K. Kim, P. Park, K. Park. Person Re-identification Method Using Text Description Through CLIP. In Proceedings of the International Conference on Electronics // Information, and Communication (ICEIC), Singapore, February 5-8. 2023. pp. 1-4.
14. F. Liu, M. Kim, Z. Ren, X. Liu. Distilling CLIP with Dual Guidance for Learning Discriminative Human Body Shape Representation // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2024. pp. 1-10.
15. Q. Han, X . He, Z. Liu, S. Liu, Y. Zhang, J. Xiang. CLIP-SCGI: Synthesized Caption-Guided Inversion for Person Re-Identification // arXiv:2410.09382. 2024
16. X. Tan, X. Gong, Y. Xiang. CLIP-based Camera-Agnostic Feature Learning for Intra- camera Person Re-Identification // arXiv:2409.19563. 2024.

VISUAL-TEXT DESCRIPTOR FOR RE-IDENTIFYING THE PERSON IN THE IMAGE

PhD stud. Savkov M.V.

Reshetnev Siberian State University of Science and Technology

Person Re-Identification (ReID) is one of the most important tasks in computer vision, particularly in the context of video surveillance and behavior analysis. Variations in viewpoint, lighting, and the appearance of the subject lead to the loss of discriminative features, reducing the effectiveness of traditional methods. This paper presents a multimodal approach based on a visual-text descriptor that combines visual data with textual descriptions of objects. Experimental results demonstrate that the proposed approach improves identification accuracy compared to state-of-the-art methods.

ИСПОЛЬЗОВАНИЕ МЕТРИКИ PIQUE ДЛЯ ОЦЕНКИ КАЧЕСТВА МЕДИЦИНСКИХ ВИДЕОИЗОБРАЖЕНИЙ

преп. Седов А.Г., доц. Хрящев В.В., проф. Приоров А.Л.

Ярославский государственный университет им. П.Г. Демидова

Рассматривается задача оценки качества видеоизображений при проведении колоноскопических исследований. Создано обучающее множество из 10 колоноскопических исследований пациентов. Произведена процедура субъективной экспертной оценки качества датасета. Предложена модификация метрики PIQUE для оценки качества видеоизображений. Вычислена корреляция между экспертными и алгоритмическими оценками качества.

Введение

Задача разработки новых алгоритмов и методов оценки качества видеоизображений сохраняет свою актуальность с учетом постоянного расширения области практического использования области анализа визуальной информации [1, 2]. Алгоритмы оценки качества делятся на эталонные и неэталонные в зависимости от наличие исходного (неискаженного) изображения – эталона. В большинстве практических случаев интерес представляет использование неэталонных алгоритмов (NR-IQA). Критерием эффективности алгоритма из класса NR-IQA является близость получаемых с его помощью оценок качества к субъективным оценкам экспертов (Mean opinion score – MOS). Для сравнения наборов оценок используется линейная или ранговая корреляция.

Одной из растущих сфер применения области цифровой обработки видеоизображений является медицина. Рассмотрим процедуру колоноскопии, которая позволяет исследовать толстую кишку и обнаруживать различные нарушения, такие как полипы, раковые заболевания, эрозии или язвенные поражения слизистой оболочки. Качество изображений в колоноскопии влияет на точность диагностики и эффективность лечения пациента. К типичным факторам, ухудшающим качество колоноскопических изображений относятся низкая освещенность, размытие вследствие движения эндоскопа, блики на поверхности слизистой оболочки кишечника [3]. В предыдущей работе мы рассмотрели эффективность известных метрик качества применительно к оценке качества колоноскопических изображений, а также обучили собственный алгоритм оценки качества на основе экспертных оценок. [4]

Для оценки качества изображения важен не один только факт наличия артефактов, но и то, какие части изображения им затронуты. Так существуют алгоритмы неэталонной оценки качества, использующие сверточные нейронные сети и обученные выявлению на изображении наиболее значимых областей. [5] Представляется что при оценке качества медицинского изображения эксперт тоже в первую очередь обращает внимание на его критически важные области. Так для колоноскопических изображений характерно наличие плохо освещенных областей, которые, как правило, находятся на дальнем плане. В этой статье мы модифицируем одну хорошо подходящую для этого метрику качества так, чтобы она отдавала приоритет при оценке хорошо освещенным областям.

База колоноскопических изображений

Было собрано множество из 1000 изображений, выбранных из 10 видеоизображений. Выбор кадров осуществлялся случайно среди множества всех отличающихся друг от друга кадров. Видеоизображения получены в разных условиях с использованием различной аппаратуры и имеют различное разрешение.

Для каждого изображения была выставлена оценка пятью экспертами. Оценка принимает значение от 1 до 5. Эксперты были проинструктированы ставить оценку «1», если на изображении ничего невозможно различить, оценку «5», если изображение имеет идеальное качество, оценку «3», если на изображении имеется область высокого качества, оценки «2» и «4» в других случаях. Был вычислен MOS – среднее арифметическое этих оценок. Мы ищем метрику качества, сильнее всего коррелирующую с MOS.

Метрика PIQUE

Были вычислены значения некоторых неэталонных метрик для множества колоноскопических изображений. Метрика PIQUE [6] показала один из лучших результатов (Таблица 1). Для вычислений использована библиотека IQA-Pytorch [7]. Звездочкой в таблице отмечены метрики, для которых меньшее значение указывает на меньшее качество изображения. В идеальном случае линейная корреляция для такой метрики составит -1

Таблица 1. Корреляция значений некоторых метрик оценки качества с MOS

Метрика	NIQE(*)	PIQUE(*)	BRISQUE(*)	PAQ-2-PIQ
Корреляция с MOS	-0.44(*)	-0.48(*)	-0.43(*)	0.25

Подобно метрике NIQE и метрике BRISQUE, для вычисления метрики PIQUE используются коэффициенты MSCN (Mean Subtracted Contrast Normalized coefficients) [8]. Для всех пикселей изображения в градациях серого вычитается локальное среднее и производится нормализация значений. Затем все изображение делится на квадратные блоки со стороной S пикселей. Если ширина или высота изображения не кратна размеру блока – края не рассматриваются.

Блоки, для которых дисперсия значений превышает эмпирически выбранный порог T_U=0.1 называются пространственно-активными. Вычисляется их количество N_{SA}. Остальные блоки далее не рассматриваются. Подразумевается, что именно пространственно-активные блоки привлекают внимание человека. Каждый пространственно-активный блок относится к одному из четырех типов: зашумленные блоки, искаженные блоки, блоки, содержащие как шум так и искажения и «хорошие» блоки (не относящиеся к трем предыдущим типам). Блокам первого и второго типа ставится оценка D, зависящая от уровня шума или искажения (Оценка тем больше, чем выше шум или искажение). Блок считается искаженным, если дисперсия в одной из прямоугольных областей на его границе оказывается ниже выбранного порога T_{STD}. Искаженным блокам свойственны артефакты блочности. Блокам третьего типа ставится оценка D=1, блокам четвертого типа присваивается оценка D=0. Итоговая метрика PIQUE затем вычисляется по формуле:

$$PIQUE = \frac{\left(\sum_{k=1}^{N_{SA}} D\right) + 1}{N_{SA} + 1}$$

Единица в числителе и знаменателе добавлена чтобы избежать численной неустойчивости. PIQUE выражает отношение количества визуально некачественных блоков среди привлекающих внимание к общему количеству блоков, привлекающих внимание. Она принимает максимальное значение 1, если все привлекающие внимание блоки некачественные и минимальное значение 0, если все привлекающие внимание блоки качественные. Значение метрики PIQUE зависит от нескольких заранее выбранных параметров. Для нашего датасета наилучшая корреляция немодифицированной оценки PIQUE с метрикой MOS составила -0.55 при следующих глобальных параметрах: S=12, T_{STD}=0.05, T_U=0.1.

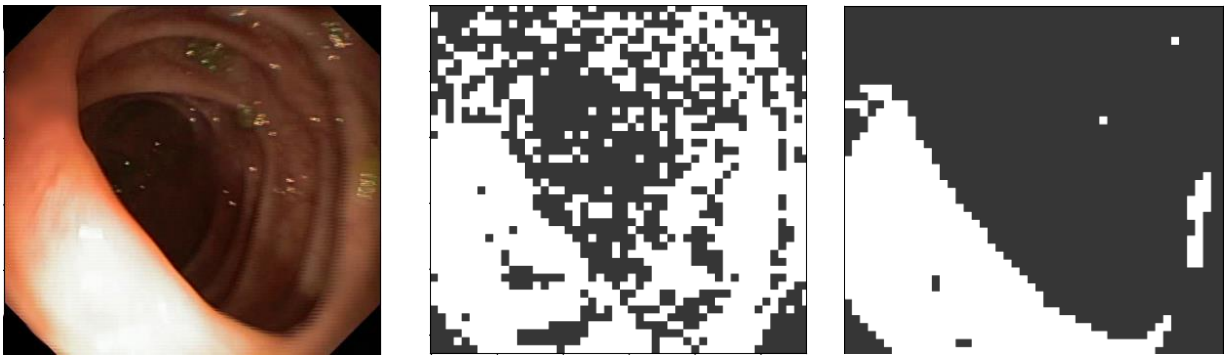
Мы добавили дополнительный критерий для определения блока как пространственно-активного. Блок считается пространственно-активным, если дисперсия значений MSCN превышает эмпирически выбранный порог T_U и если средняя яркость пикселей в блоке *исходного изображения* превышает 100. (Яркость пикселя принимает значения от 0 до 255).

После добавления дополнительного критерия корреляция с MOS улучшилась до -0.7.

Заключение

Рассмотренная метрика PIQUE и ее модификация обладают следующими преимуществами:

- простая распараллеливаемость, обусловленная ориентированностью на работу с блоками;
- высокая скорость работы;
- алгоритм может помимо самой метрики указать на зашумленные, искаженные и активные блоки, что позволяет визуально оценить эффективность его работы (Рисунок 1).



(а) (б) (в)

Рисунок 1. Колоноскопическое изображение и созданные модифицированным алгоритмом PIQUE маски: исходное изображение (а), блоки с шумом (б), блоки пространственной активности (в)

Совершенствование методов выявления области интереса как первичный шаг при оценке качества изображения видится перспективным направлением дальнейшей работы.

Литература

1. Дворкович В.П., Дворкович А.В. Цифровые видеоинформационные системы (теория и практика). Техносфера, 2012. 1008 с.
2. Приоров А.Л., Хрящев В.В., Топников А.И. Обработка и передача мультимедийной информации: учебное пособие. Ярославль: ЯрГУ. 2022. 248 с.
3. Палевская С.А., Короткевич А.Г. Эндоскопия желудочно-кишечного тракта. М.: ГЭОТАР-Медиа, 2020. 752 с.
4. Хрящев В.В., Седов А.Г., Приоров А.Л. Неэталонная оценка качества изображений колоноскопических исследований // Цифровая обработка сигналов. 2024. №3. С. 44-49.
5. Chen C., Mo J., Hou J., Wu H., Liao L., Sun W., Yan Q., Lin W. TOPIQ: A Top-Down Approach From Semantics to Distortions for Image Quality Assessment // IEEE Transactions on Image Processing, vol. 33, pp. 2404-2418, 2024.
6. Venkatanath N, Praneeth D, Maruthi Chandrasekhar Bh, S. S. Channappayya and S. S. Medasani, "Blind image quality evaluation using perception based features," 2015 Twenty First National Conference on Communications (NCC), Mumbai, India, 2015, pp. 1-6.
7. Chaofeng Chen, Jiadi Mo. IQA-PyTorch: PyTorch Toolbox for Image Quality Assessment // https://github.com/chaofengc/IQA-PyTorch, 2022.
8. A. Mittal, A. K. Moorthy, and A. C. Bovik, “Noreference image quality assessment in the spatial domain,” IEEE Transactions on Image Processing, vol. 21, no. 12, pp. 4695-4708, 2012.

USING THE PIQUE METRIC TO ASSESS THE QUALITY OF MEDICAL VIDEO IMAGES

Lecturer A.G. Sedov, Assoc. Prof. V.V. Khryashchev, Prof. A.L. Priorov

P.G. Demidov Yaroslavl State University

The problem of assessing the quality of video images during colonoscopic examinations is considered. A training set of 10 colonoscopic examinations of patients was created. A procedure for subjective expert assessment of the quality of the dataset was performed. A modification of the PIQUE metric for assessing the quality of video images is proposed. The correlation between expert and algorithmic quality assessments is calculated.

СРАВНИТЕЛЬНЫЙ АНАЛИЗ ИСПОЛЬЗОВАНИЯ КЛАССИЧЕСКИХ И НЕЙРОСЕТЕВЫХ АЛГОРИТМОВ ДЛЯ ФИЛЬТРАЦИИ СПЕКЛ-ШУМА ИЗ ИЗОБРАЖЕНИЙ

асп. Сенников А.В., асп. Ларионов Р.В., к.т.н., доц. Хрящев В.В.

Ярославский государственный университет им. П.Г. Демидова

В работе рассматриваются способы фильтрации спекл-шума на изображениях. Используются такие фильтры, как фильтр Ли, билатеральный и медианный фильтр, фильтры BM3D и DWM, sar_transformer. Для оценки качества выбраны метрики MSE, PSNR, SSIM, BRISQUE, ENL и SMPI. Тестирование производится на стандартном восьмибитном изображении «Лодка». Для моделирования спекл-шума используется мультипликативная модель шума с нормальным распределением с единичным математическим ожиданием и различными значениями дисперсии. Для сильной степени зашумленности, выраженной значением PSNR в 16.11 дБ, больше всего выделяются фильтр BM3D и нейросетевая модель sar_transformer. Изображение после фильтра BM3D ближе всего визуально к исходному изображению, а sar_transformer лучше всего удаляет непосредственно спекл-шум, но сильно сглаживает изображение.

Введение

В настоящее время использование данных дистанционного зондирования земли нашло широкий спектр применения. Задачей такого направления, например, является оценка лесной биомассы, управление сельскохозяйственными и природными ресурсами, геология, предупреждение стихийных бедствий [1]. Данные со спутника в оптическом диапазоне, полученные пассивными приемниками, не самый подходящий источник информации в подобных задачах, так как сильно зависят от погодных условий. Поэтому для такого рода задач используют спутниковые изображения в радиодиапазоне, полученные активными радиолокационными устройствами такими, как радары с синтезированной апертурой (SAR, РСА). Благодаря ним можно проводить исследования, игнорируя атмосферные условия [2-4]. Однако при это на изображениях, получаемых с помощью систем РСА, присутствует высокий уровень шума, обусловленного природой обратного рассеяния радара. Самым большим источником шума на изображении РСА является спекл-шум, который вызывается взаимодействием когерентного луча радара и снимаемой относительно шероховатой поверхности [5].

Целью данной работы является исследование алгоритмов (классических, современных нейросетевых) удаления спекл-шума с цифровых изображений. При разработке фильтров для спекл-шума на изображениях РСА большой проблемой является то, что изображений со спутника изначально без спекл-шума нет. Данная проблема затрудняет качественную оценку работы фильтров. В связи с этим в работе оценка фильтрации производится с помощью стандартных тестовых изображений, на котором изначально спекл-шум отсутствует. Проведенное исследование позволит выработать рекомендации по использованию алгоритмов для фильтрации спекл-шума в реальных практических задачах обработки изображений дистанционного зондирования земной поверхности, полученных в радиодиапазоне.

Модель спекл-шума

Спекл-шум проявляется как зернистый или пятнистый паттерн на изображениях, что затрудняет их интерпретацию и анализ. Природа спекл-шума заключается в том, что радиолокационные системы отправляют когерентные микроволновые сигналы, которые отражаются от поверхности Земли и возвращаются к приемнику, после чего эти отраженные волны складываются когерентным образом, что приводит к интерференционным паттернам. В результате интерференции происходит случайное усиление или ослабление сигнала, создавая зернистую структуру. Для моделирования используется мультипликативная модель шума, в которой исходный сигнал умножается на случайную величину с единичным математическим ожиданием [6, 7]. Эта величина может иметь распределение Релея, Гамма распределение или нормальное распределение. В данной работе используется нормальное распределение. Математически модель можно записать следующим образом:

$$noised_signal(x) = signal(x) * S,$$

где $noised_signal$ – зашумленный сигнал, $signal$ – исходный сигнал, S – случайное распределение.

Фильтрация спекл-шума

Возможность эффективного удаления спекл-шума с изображений имеет решающее значение в задачах ДЗЗ. Этой теме посвящено много работ. Для устранения спекл-шума классическим считается использование пространственных фильтров. В данной работе используется такие алгоритмы как фильтр Ли, билатеральный и медианный фильтр, BM3D [8] и DWM [9].

Также в задаче фильтрации спекл-шума в современных работах используются алгоритмы глубокого обучения. В их случае большой проблемой является нехватка эталонных изображений без спекл-шума для обучения. В данной работе для сравнения рассматривается алгоритм sar_transformer [10]. В источнике предлагается модель типа энкодер-декодер. Согласно авторам, данная модель работает лучше, чем другие способы фильтрации.

Результаты моделирования

Для исследования алгоритмов фильтрации было использовано стандартное тестовое восьмибитное изображение размером 512 на 512 пикселей – «Лодка». Было принято использовать именно такое изображение, а не изображение РСА, в связи с отсутствием изображений РСА без спекл-шума. Оно было зашумлено в различной степени с использованием вышеописанной модели спекл-шума с нормальным распределением. Степень зашумления задавалась дисперсией нормального распределения, которая принимала значения 0,075, 0,1, 0,2, 0,3, 0,4. Математическое ожидание всегда было равно единице, чтобы не изменялась яркость изображения. Каждое полученное изображение было отфильтровано вышеописанными фильтрами. Для каждого зашумленного и отфильтрованного изображения были измерены метрики MSE, PSNR, SSIM, BRISQUE, ENL и SMPI [11]. В таблице 1 представлены значения метрик, полученных при тестировании. В ней var – дисперсия, lee – фильтр Ли, bilateral – билатеральный фильтр, median – медианный фильтр. Данная таблица содержит значения метрик для слабой и сильной степени зашумленности, имеющей характерные значения PSNR в 27,85 дБ и 16,11 дБ соответственно и выраженной значениями дисперсии 0,075 и 0,3.

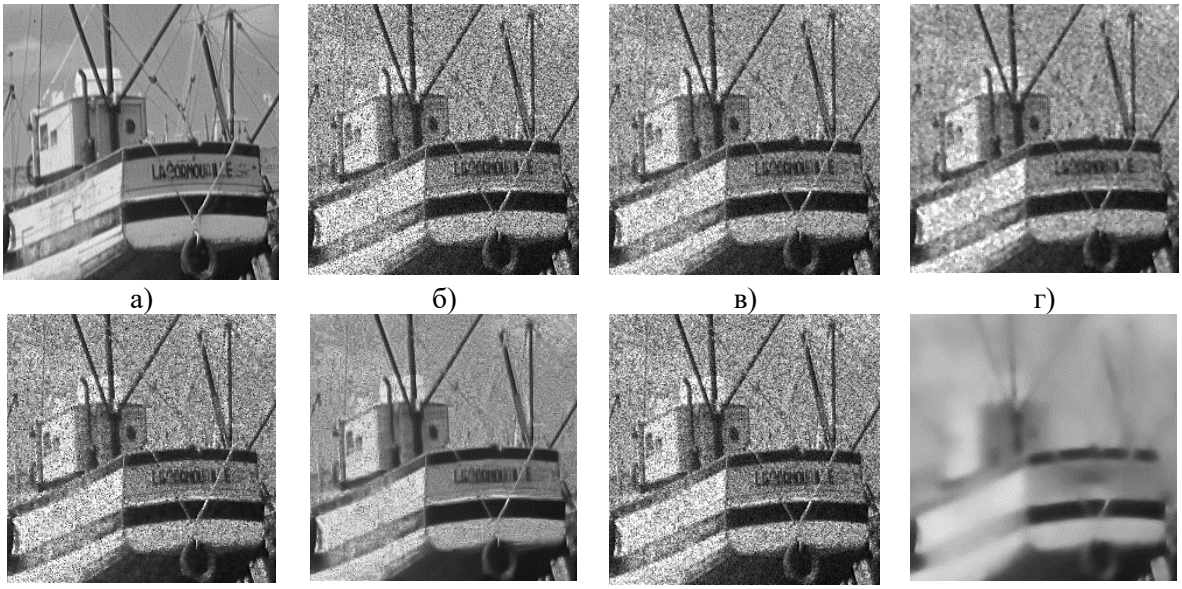
Таблица 1. Значения полученных метрик оценки качества

var	метрика	+ шум	lee	bilateral	median	bm3d	dwm	sar_transformer
0,075	mse	106,78	72,52	39,68	98,42	69,69	107,01	609,15
0,075	psnr	27,85	29,53	31,94	28,20	29,10	27,84	19,31
0,075	ssim	0,72	0,77	0,85	0,79	0,76	0,72	0,49
0,075	enl	51,14	60,77	92,90	90,29	172,02	51,14	823,67
0,075	smpi	1,74	1,69	1,31	1,54	1,11	1,75	6,05
0,075	brisque	157,03	145,27	155,69	152,57	124,23	157,03	114,79
0,3	mse	1591,70	992,79	1172,34	378,10	290,82	1418,52	635,89
0,3	psnr	16,11	18,16	17,44	22,35	23,49	16,61	18,97
0,3	ssim	0,28	0,34	0,35	0,48	0,51	0,29	0,48
0,3	enl	9,67	14,99	13,22	38,31	52,88	10,66	932,17
0,3	smpi	7,25	5,98	6,18	4,43	3,39	7,02	6,42
0,3	brisque	171,46	133,21	139,08	169,95	146,95	171,69	112,60

Для слабой зашумленности по метрикам PSNR и SSIM наибольшее улучшение дает билатеральный фильтр, имеющий PSNR в 31,94 дБ и SSIM 0,85. Для сильной зашумленности по данным метрикам лучше работает фильтр BM3D, значительно повышающий PSNR с 16.11 дБ до 23,49 дБ. По метрике SMPI фильтр BM3D также имеет наибольшее улучшение для обеих степеней зашумленности. Нейросетевой алгоритм sar_transformer для слабой зашумленности только ухудшает изображение. Для сильной зашумленности улучшает, но в меньшей степени, чем Bm3D. При этом по метрике ENL sar_transformer имеет наибольший прирост, превышающий другие на порядок. Это означает, что он гораздо сильнее сглаживает изображение.

Визуальное сравнение методов фильтрации представлено на рисунке 1. Оно подтверждает значения метрик. По нему видно, что наиболее похожим на исходное изображение является

изображение после обработки фильтром BM3D. При этом sar_transformer полностью удаляет спекл-шум, но сильно сглаживает, что приводит к тому, что теряется очень много деталей.



а) б) в) г)

д) е) ж) З)

Рисунок 1 – Результаты тестирования на изображении «Лодка»: а) исходное изображение, б) изображение с шумом при var = 0,3 и обработка фильтрами в) lee, г) median, д) bilateral, е) Bm3d, ж) DWM, з) sar_transformer

Заключение

В рамках данного исследования были рассмотрены методы удаления спекл-шума, его модель и метрики для оценки качества его фильтрации. В качестве методов удаления шума были выбраны фильтр Ли, BM3D и DWM, билатеральный и медианный фильтр, а также нейросетевая модель sar_transformer. Для моделирования спекл-шума использовалась мультипликативная модель шума с нормальным распределением. Для оценки качества были использованы метрики MSE, PSNR, SSIM, BRISQUE, ENL и SMPI. Исследование на изображении «Лодка» показало, что как визуально, так и по метрикам лучше отрабатывает фильтр BM3D. Он повышает PSNR с 16.11 дБ до 23,49 дБ и имеет наибольшее улучшение метрики SMPI. Нейросетевой фильтр sar_transformer полностью удаляет спекл-шум с изображения, о чем свидетельствует увеличение метрики ENL до 932,17. Отрицательной стороной его работы является сильное сглаживание изображения, приводящее к потере основных деталей.

Исследование выполнено при финансовой поддержке Правительства Ярославской области в рамках научного проекта № 16НП/2024.

Литература

1. Рис У.Г. Основы дистанционного зондирования / М.: Техносфера, 2006. 336 с.

2. Гук А.П., Евстратова Л.Г. Дистанционное зондирование и мониторинг территорий / М.: КУРС, 2025. 224 с.

3. Ларионов Р. В., Сенников А. В., Хрящев В. В. Разработка алгоритма определения изменений лесного покрова на изображениях со спутника Sentinel-2 //33-я Международная конференция по компьютерной графике и машинному зрению «ГрафиКон2023», Москва, ИПУ РАН, 2023. С.721-726. – 2023. – Т. 33. – С. 721-726.

4. Larionov R., Sennikov A., Khryashchev V. Water Segmentation from SAR Images with the Presence of Speckle Noise //The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. – 2024. – Т. 48. – P. 95-100.

5. Maity A. et al. A comparative study on approaches to speckle noise reduction in images //2015 International Conference on Computational Intelligence and Networks. – IEEE, 2015. – P. 148-155.

6. Волков, В. Ю. Моделирование спекл-изображений и селекция объектов / СПбНТОРЭС: труды ежегодной НТК. – 2021. – № 1(76). – С. 71-74.

7. Singh P., Pandey R. Speckle noise: Modelling and implementation // Intern. J. of Circuit Theory and Applications. – 2016. – Vol. 9, iss. 17. – P. 8717–8727.

8. Yahya A. A. и др. BM3D image denoising algorithm based on an adaptive filtering // Multimedia Tools and Applications. – 2020. – Т. 79. № 27. – P. 20391–20427.

9. Dong Y., Xu S. A New Directional Weighted Median Filter for Removal of Random-Valued Impulse Noise // IEEE Signal Processing Letters. – 2007. – Т. 14. № 3. – P. 193–196.

10. Perera M. V. et al. Transformer-based SAR image despeckling //IGARSS 2022-2022 IEEE International Geoscience and Remote Sensing Symposium. – IEEE, 2022. – P. 751-754.

11. Wang X., Ge L., Li X. Evaluation of filters for ENVISAT ASAR speckle suppression in pasture area //ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences. – 2012. – Т. 1. – P. 341-346.

COMPARATIVE ANALYSIS OF THE USE OF CLASSICAL AND NEURAL NETWORK ALGORITHMS FOR DESPECKLING IMAGES

Sennikov A.V., Larionov R.V., Khryashchev V.V.

P.G. Demidov Yaroslavl State University

The paper discusses ways to despekle images. Filters such as Lee filter, bilateral and median filters, BM3D and DWM filters, and sar_transformer are used. The metrics MSE, PSNR, SSIM, BRISQUE, ENL, and SMPI were selected for quality assessment. Evaluating is performed on a standard 8-bit "Boat" image. To model speckle noise, a multiplicative noise model with a normal distribution with a constant mean and different values of variance is used. For a high degree of noise, expressed by a PSNR value of 16.11 dB, the BM3D filter and the sar_transformer neural network model stand out the most. The image after the BM3D filter is closest visually to the original image, and sar_transformer is best at removing speckle noise directly, but greatly smoothing the image.

ИССЛЕДОВАНИЕ АЛГОРИТМОВ АВТОМАТИЧЕСКОЙ КЛАССИФИКАЦИИ ВИДОВ ЦИФРОВОЙ МОДУЛЯЦИИ РАДИОСИГНАЛОВ ПРИ ПОМОЩИ СВЕРТОЧНЫХ НЕЙРОННЫХ СЕТЕЙ

асп. Ушакова А.М.

Ярославский государственный университет им. П.Г. Демидова

Исследуется применение сверточных нейронных сетей (СНС) для автоматической классификации модулированных радиосигналов, что является важной задачей в области радиосвязи и обработки сигналов. Используются данные из базы модулированных сигналов DeepSig, а именно набор RadioML2018.01A, который включает в себя различные типы аналоговой и цифровой модуляции. В ходе исследования разработана модель на основе СНС ResNet50. Представлены графики, иллюстрирующие зависимости точности обучения сети от числа эпох, а также матрица ошибок классификации, демонстрирующая результаты работы модели для различных типов модулированных радиосигналов.

Введение

В последние десятилетия наблюдается значительный рост интереса к автоматическому распознаванию (классификации) видов цифровой модуляции, что связано с увеличением объема данных, передаваемых по беспроводным сетям, и необходимостью эффективного использования радиочастотного спектра. Современные технологии связи, такие как 4G и 5G, требуют высокой плотности передачи данных и надежности, что делает важным развитие методов, позволяющих быстро и точно идентифицировать используемые виды модуляции. Это, в свою очередь, способствует оптимизации работы беспроводных сетей передачи данных и улучшению качества обслуживания пользователей.

Цифровая модуляция (манипуляция) является значимым элементом в современных системах беспроводной связи, так как она в значительной степени определяет способ передачи информации через радиоканал. Разнообразие видов цифровой модуляции требует от систем радиомониторинга и анализа способности к быстрой и точной идентификации используемых систем передачи информации. Традиционные методы анализа, основанные на ручных алгоритмах и экспертных системах, часто оказываются недостаточно эффективными для обработки больших объемов данных в реальном времени [1-2].

Сверточные нейронные сети, являющиеся представителем технологий глубокого машинного обучения, представляют собой мощный инструмент для автоматического распознавания видов цифровой модуляции. Благодаря своей способности извлекать сложные признаки из изображений и временных рядов, они могут значительно улучшить точность классификации модуляции, что делает их удобными для применения в задачах распознавания сигналов.

Кроме того, вторичное использование радиочастотного спектра становится все более актуальным в условиях растущего спроса на беспроводные услуги. В то время как спектр в области радиочастот становится все более загруженным, эффективное управление и перераспределение частот могут привести к значительному улучшению качества связи и уменьшению помех. Автоматическое распознавание видов модуляции играет важную роль в этом процессе, позволяя операторам адаптироваться к изменениям в спектре и эффективно использовать доступные ресурсы [3].

Таким образом, работа посвящена исследованию возможностей применения сверточных нейронных сетей для автоматического распознавания видов цифровой модуляции, в том числе в условиях низкого отношения сигнал/шум.

Экспериментальная часть

Для исследования взята одна из доступных в открытом информационном пространстве база данных RADIOML2018.01.A. Хранящиеся в формате HDF5 [4] данные для нее сгенерированы в том числе и в реальных условиях, с учетом шума и помех, что делает исследование более близким к практическим задачам.

Данный набор содержит 2 555 904 кадра, каждый из которых представляет собой 1024 выборки сложных временных рядов. Эти выборки состоят из синфазных (I) и квадратурных (Q) составляющих с плавающей запятой, представленных в формате (1024, 2). Данные

сгенерированы для 26 уровней отношения сигнал/шум (ОСШ) от -20 дБ до +30 дБ с шагом 2 дБ. Для каждой комбинации вида цифровой модуляции и ОСШ создано 4096 кадров. На рисунке 1 представлено описание базы данных RADIOML2018.01.A.

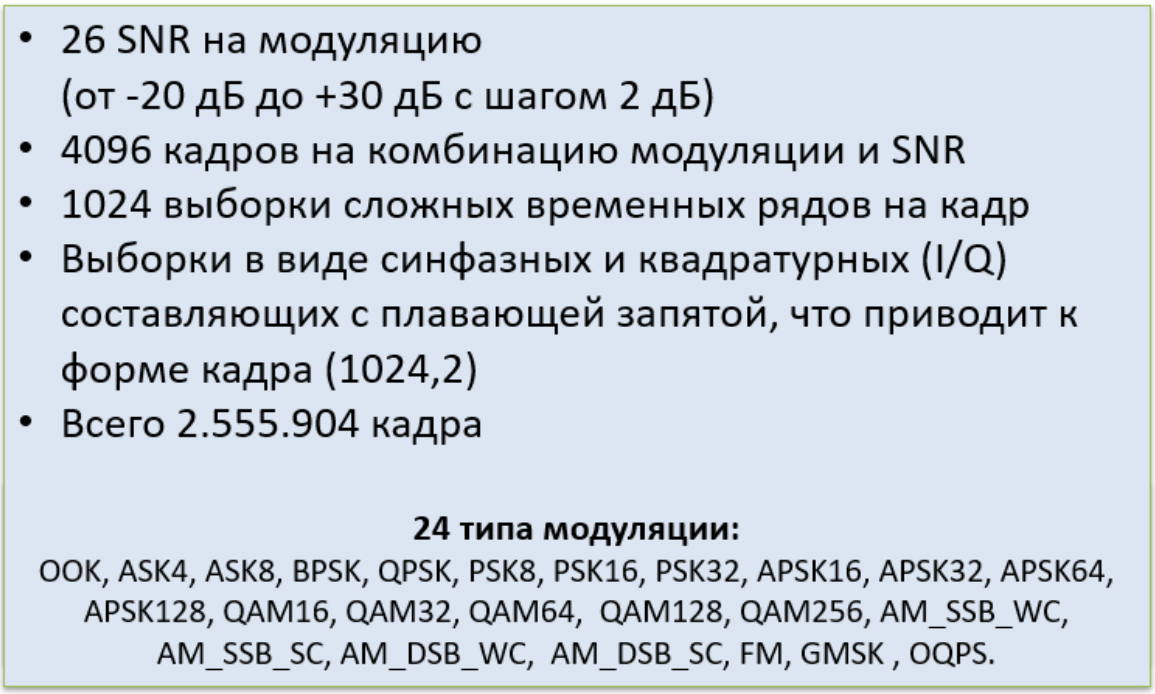


Рисунок 1 – Описание набора данных RADIOML2018.01.A

Каждому кадру соответствуют три группы данных в формате HDF5: X – компоненты IQ кадра (массив (1024, 2)) представляет собой список, включающий в себя составляющие комплексного числа, т.е. отсчёты сигнала, а именно квадратуры; Y – тип модуляции (одно значение) – это список, в котором из 24 элементов только один имеет значение 1, а остальные – 0, следовательно, данный список является классовой меткой сигнала; Z – уровень ОСШ (одно значение) – список с одним элементом (целым числом).

Всего в базе представлены 24 вида модуляции, включая как классические, так и более современные: OOK, 4-ASK, 8-ASK, BPSK, QPSK, 8-PSK, 16-PSK, 32-PSK, 16-APSK, 32-APSK, 64-APSK, 128-APSK, 16-QAM, 32-QAM, 64-QAM, 128-QAM, QAM256, AM_SSB_WC, AM_SSB_SC, AM_DSB_WC, AM_DSB_ SC, FM, GMSK, OQPS.

Из данных видов модуляции в исследовании рассматриваются только цифровые модуляции, широко используемые в современных системах радиосвязи: BPSK, QPSK, 16QAM, 64QAM, 128QAM и 256QAM с ОСШ, имеющим фиксированное значение, равное 0 дБ. Количество сигналов для тренировочного набора – 60000, для валидационного набора – 12000, для тестового набора – также 12000 сигналов.

Количество отсчетов, используемых для представления одного символа цифровой модуляции, постоянно. Спектры сигналов в обучающей выборке сосредоточены в области нулевой частоты.

Обучение нейронной сети

Для обучения СНС размер мини-пакета установлен равным 64. Количество эпох обучения сверточной нейронной сети ResNet50 составляет 30. В процессе исследования длительность символов (временной интервал, занимаемый одним символом) оставалась постоянной.

В исследовании рассматривается только один тип искажений — аддитивный белый гауссовский шум (АБГШ). Это наиболее распространенная модель шума, которая предполагает, что шум добавляется к сигналу, имеет равномерную спектральную плотность (белый шум) и распределен по гауссовскому (нормальному) закону.

Архитектура ResNet50 начинается с входного слоя, принимающего исходные данные. Далее идет первый слой, где выполняется свертка с размером ядра 7x7 и шагом 2, за которой следует батч-нормализация и функция активации ReLU. Затем идет слой подвыборки (max pooling) с размером 3x3 и шагом 2.

Основная часть архитектуры состоит из 16 остаточных блоков, сгруппированных в 4 основных слоя. Первый слой включает 3 остаточных блока с 64 фильтрами, второй слой – 4 блока с 128 фильтрами, третий слой – 6 блоков с 256 фильтрами, а четвертый слой – 3 блока с 512 фильтрами. Каждый остаточный блок состоит из двух сверток 3x3, за которыми следуют батч-нормализация и активация ReLU. В некоторых блоках используется изменение размерности (downsampling) для согласования размеров входа и выхода [5-7].

После всех остаточных блоков применяется глобальное усреднение (global average pooling), что позволяет уменьшить размерность и подготовить данные для классификации. Финальный слой представляет собой полносвязный слой, который использует функцию активации Softmax для классификации изображений на заданное количество классов (например, 1000 для базы изображений ImageNet). На рисунке 2 представлено описание архитектуры сети ResNet50.

Название слоя	ResNet50
conv1	7x7, 64, stride 2
pool1	3x3, max pool stride 2
conv2_x	$\begin{bmatrix} 1x1, & 64 \\ 3x3, & 64 \\ 1x1, & 256 \end{bmatrix}$ x 3
conv3_x	$\begin{bmatrix} 1x1, & 128 \\ 3x3, & 128 \\ 1x1, & 512 \end{bmatrix}$ x 4
conv4_x	$\begin{bmatrix} 1x1, & 256 \\ 3x3, & 256 \\ 1x1, & 1024 \end{bmatrix}$ x 6
conv5_x	$\begin{bmatrix} 1x1, & 512 \\ 3x3, & 512 \\ 1x1, & 2048 \end{bmatrix}$ x 3
fc1	9x1, 2048, stride 1
pool_time	1xN, avg pool stride 1

Рисунок 2 – Описание архитектура СНС ResNet50

На рисунке 3 представлены зависимости точности классификации указанных типов цифровой модуляции сети ResNet50 и потерь от количества эпох.

В начале обучения наблюдается рост точности классификации, что связано с тем, что модель быстро обучается распознавать основные закономерности в данных. По мере увеличения количества эпох рост точности замедляется, что свидетельствует о постепенном приближении модели к оптимальным параметрам.

График потерь демонстрирует, насколько хорошо модель минимизирует ошибку между предсказанными и истинными значениями.

С каждой эпохой потери уменьшаются, что свидетельствует о том, что модель улучшает свои предсказания. Если значения потерь перестают уменьшаться или начинают колебаться, это может указывать на переобучение модели или на необходимость корректировки гиперпараметров, таких как скорость обучения.

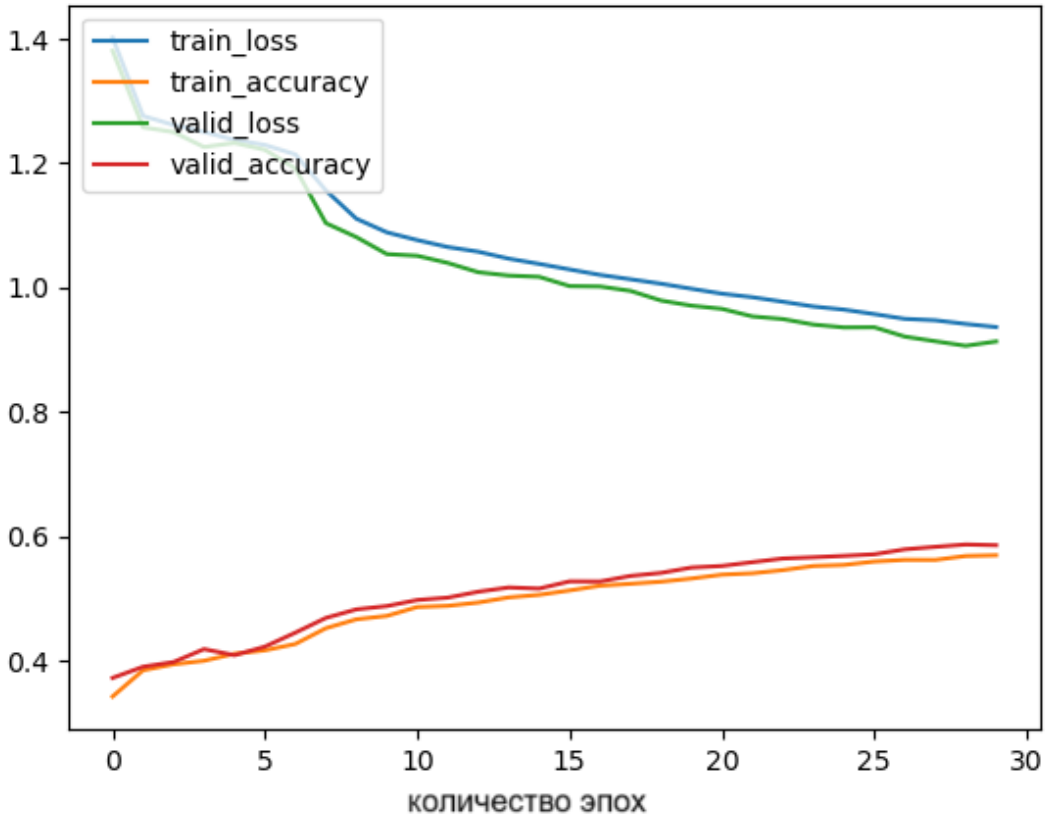


Рисунок 3 – Зависимость точности классификации и потерь от количества эпох

С увеличением количества эпох точность для валидационных данных не опускается ниже производительности тренировочного набора, из этого следует, что отсутствует переобучение нейросети.

Установлено, что модель способна правильно классифицировать виды цифровой модуляции со средней точностью распознавания 61,88%.

На рисунке 4 представлена матрица ошибок классификации видов модуляции, построенная для тестовых данных, которая позволяет визуализировать качество классификации модели. Она показывает, насколько точно нейросетевая модель идентифицирует каждый из классов, сравнивая предсказанные значения с истинными.

Из матрицы ошибок выявлено, что наилучшее распознавание среди всех видов цифровой модуляции демонстрирует BPSK с вероятностью правильной классификации 92,19%, что свидетельствует о высокой надежности модели для этого типа сигналов. Наихудший результат наблюдается для модуляции 256-QAM, где вероятность правильного распознавания составляет всего 35,23%, что может быть связано с высокой сложностью и схожестью характеристик этого типа модуляции с другими подобными, особенно в условиях действия шума.

Модель также показывает хорошие результаты для QPSK (79,72%), 16-QAM (70,22%), 64-QAM (51,65%) и 128-QAM (42,24%), однако с увеличением сложности модуляции точность классификации снижается.

Анализ матрицы ошибок также позволяет выявить, что модель практически не путает такие типы модуляции, как BPSK и 64-QAM, с вероятностью ложного распознавания всего 0,23%. Это говорит о том, что данные виды модуляции имеют существенные различия в своих характеристиках, которые модель успешно выделяет. В то же время наибольшая путаница возникает между модуляциями 64-QAM и 256-QAM, где вероятность ложного распознавания достигает 23,98%. Это может быть связано с тем, что эти типы модуляции имеют схожие спектральные и амплитудные характеристики, особенно в условиях действия искажений и шума, что затрудняет их различение.

Матрица ошибок классификации видов модуляции не только подтверждает общую эффективность модели, но и выявляет области, где требуется существенная доработка,

например, для улучшения классификации высокоуровневых видов цифровых модуляций, таких как 256-QAM.

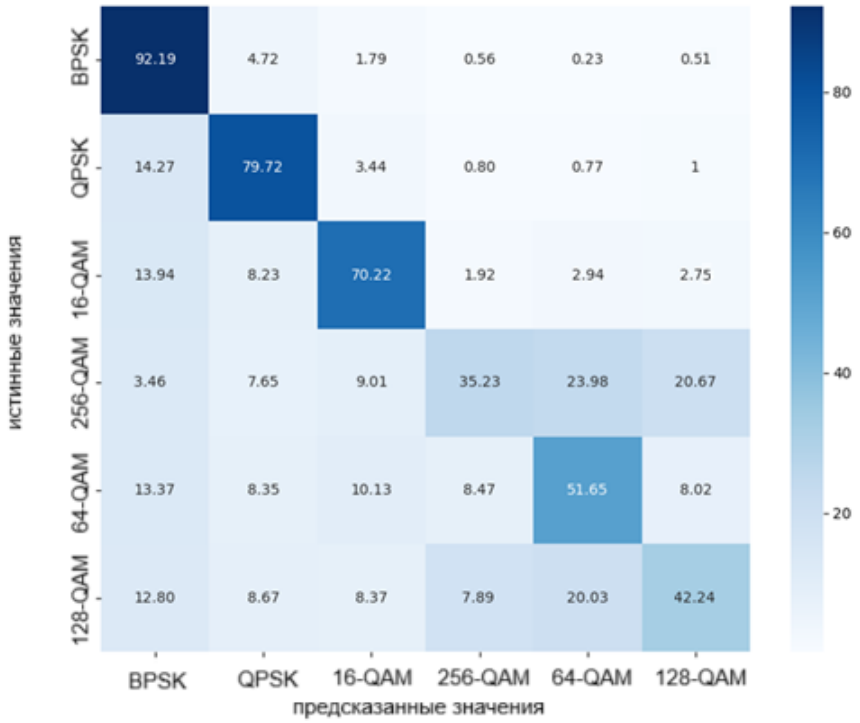


Рисунок 4 – Матрица ошибок классификации видов модуляции

Заключение

Проведено исследование алгоритма автоматической классификации видов цифровой модуляции радиосигналов при помощи сверточной нейронной сети ResNet50, обученной и протестированной с использованием базы данных RADIOML2018.01.A, где радиосигналы хранятся в формате HDF5.

Полученные результаты не являются сравнением существующих классификаторов видов цифровой модуляции, основанных на лучших экспертных характеристиках, однако они демонстрируют, что для более простых и соответственно более помехоустойчивых видов модуляции, таких как BPSK и QPSK, обученная модель представляет собой мощный инструмент классификации указанных видов цифровой модуляции. Средняя точность распознавания вида цифровой модуляции составила 61,88% при достаточно низком ОСШ, равном 0 дБ.

Применение данной модели актуально в таких системах, где необходима высокоточная классификация видов цифровой модуляции с низким отношением сигнал/шум [8-10]. К ним, например, относятся системы радиомониторинга, которые постоянно расширяют свой аппаратный и программный потенциал в плане анализа текущей радиообстановки на подответственной им территории.

Полученные результаты свидетельствуют о перспективности применения глубоких нейронных сетей, в частности архитектуры ResNet50, для задач классификации видов цифровой модуляции в сложных условиях работы радиосистемы передачи данных. Это подтверждает возможность использования рассмотренных методов в системах, где критически важна высокая точность распознавания при ограниченных энергетических ресурсах сигнала. Дальнейшие исследования могут быть сосредоточены на оптимизации гиперпараметров нейросетевой модели, интеграции дополнительных методов предобработки данных, таких как спектральное преобразование или фильтрация, а также на изучении влияния различных типов зашумления на качество классификации.

Кроме того, перспективным направлением является адаптация разработанной нейросетевой модели для работы с более широким спектром видов модуляции, включая сложные схемы,

такие как OFDM или QAM высокой кратности, что позволит расширить область применения алгоритма в современных системах связи и радиомониторинга, где указанные схемы достаточно широко используются.

Исследование выполнено при финансовой поддержке Правительства Ярославской области в рамках научного проекта от 25 декабря 2024 г. №13НП/2024.

Литература

1. Zhang D., Ding W., Zhang B., Xie C., Li H., Liu C., Han J. Automatic modulation classification based on deep learning for unmanned aerial vehicles // MDPI Sensors (Basel). 2018.
2. Zhu Z., Nandi A.K. Automatic Modulation Classification Principles // Algorithms and Applications. 2015.
3. Аджемов С.С., Чиров Д.С., Терешонок М.В. Распознавание видов цифровой модуляции в системах когнитивного радио. М.: МТУСИ, 2018. 224 с.
4. RF datasets for machine learning [Электронный ресурс]. Режим доступа: https://www.deepsig.ai/datasets.
5. Sainath T.N., Vinyals O., Senior A., Sak H. Convolutional long short-term memory, fully connected deep neural networks // IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). 2015. pp. 4580-4584.
6. Лекун Я. Как учится машина: Революция в области нейронных сетей и глубокого обучения / пер. с фр. М.: Альпина, 2021. 335 с.
7. Хайкин С. Нейронные сети: полный курс / 2-е издание. М.: Вильямс, 2008. 1104 с.
8. Lv H., Zhou X., Huo J., Yuan J. Joint OSNR monitoring and modulation format identification on signal amplitude histograms using convolutional neural network // Optical Fiber Technology. Volume 61. January 2021.
9. Mishkina D., Sergievskiy N., Matas J. Systematic evaluation of convolution neural network advances on the Imagenet // Computer Vision and Image Understanding. 2017. vol. 161. pp. 11-19.
10. Blossom E. Gnu radio: tools for exploring the radio frequency spectrum // Linux journal. vol. 2004, no. 122, p. 4, 2004.

INVESTIGATION OF ALGORITHMS FOR AUTOMATIC CLASSIFICATION OF TYPES OF DIGITAL MODULATION OF RADIO SIGNALS USING CONVOLUTIONAL NEURAL NETWORKS

Ushakova A.M.

P.G. Demidov Yaroslavl State University

The application of convolutional neural networks (SNN) for the automatic classification of modulated radio signals is being investigated, which is an important task in the field of radio communications and signal processing. Data from the Deepsign modulated signal database is used, namely the RadioML2018.01A set, which includes various types of analog and digital modulation. In the course of the research, a model based on the ResNet50 SNA was developed. Graphs illustrating the dependence of network learning accuracy on the number of epochs are presented, as well as a classification error matrix demonstrating the results of the model for various types of modulated radio signals.

ОЦЕНКА ШУМОВЫХ ХАРАКТЕРИСТИК ЦИФРОВЫХ КАМЕР ПО ОДНОМУ ИЗОБРАЖЕНИЮ

асп. Козлов А.В.[1,2]

[1] «Национальный исследовательский ядерный университет «МИФИ»
[2] ФГУП Всероссийский научно-исследовательский институт физико-технических и радиотехнических измерений

Представлен алгоритм оценки шумовых параметров фотосенсоров цифровых камер из данных, полученных в результате применения алгоритма оценки шумов одного изображения. Проведена апробация предложенного метода на снимках подготовленной сцены и на снимке произвольной бытовой сцены. Получено, что предложенный метод даёт более точные оценки по сравнению с прямой аппроксимацией данных, получаемых в результате применения алгоритма оценки шумов одного изображения выражением для суммарного среднеквадратического отклонения всех составляющих шума.

Введение

На сегодняшний день цифровые регистраторы используются в широком спектре задач. Для каждой задачи важны определённые характеристики камеры. Среди всех характеристик цифровых камер менее всего производители сенсоров уделяют внимание шумам. В современных условиях, при решении инновационных задач данная характеристика становится не менее важной, чем остальные. Чтобы обеспечить быстрые и надёжные измерения шумов цифровых камер разработано множество методов, для реализации которых требуется различное число снимков. Наиболее точный метод был предложен европейской ассоциацией машинного зрения (European Machine Vision Association - EMVA), однако его реализация требует вложения труда и времени. Поэтому для упрощения и ускорения оценки шумовых характеристик камеры были разработаны методы на основе автоматической сегментации. Временные затраты и трудоёмкость при этом существенно снизились, однако оценка шумов, как и в EMVA требует снимков определённой сцены, что означает необходимость подготовки специальных условий.

Идеальным решением задачи оценки шумов цифровых камер является алгоритм, позволяющий по кадру произвольной сцены оценить все средние шумовые характеристики для сенсора по отдельности. Существуют алгоритмы, оценивающие шумы по одному кадру. Однако, важно отметить, что такие алгоритмы предназначены для оценки шума изображения, а не камеры, на которую данное изображение было зарегистрировано. Поскольку итогом работы таких алгоритмов оценки шумов изображения является зависимость отклонения от уровня сигнала, разделение параметров шума представляет из себя отдельную задачу, которой и посвящена данная работа.

Шумовые параметры фотосенсоров цифровых камер

Для описания шумов цифровых камер в данной работе используется модель, представленная в EMVA 1288 [1]. Шумы цифровых камер могут в общем быть разделены на две составляющие: случайные и детерминированные. Случайные шумы меняются от одного снимка к другому, детерминированные же остаются постоянными. К первому типу относится дробовой шум, возникающий вследствие накопления различного числа фотоэлектронов в ячейках пикселя во время освещения сенсора. Кроме того, первый тип включает шум, порождаемый их случайным рождением и темновыми токами в отсутствии освещения. Дробовой шум также называют световым временным или Пуассоновским. Он возникает при подаче сигнала на фотосенсор, является случайным и хорошо описывается статистикой Пуассона. Случайная составляющая шума, существующая и в отсутствии сигнала называется темновой временной соответственно. Детерминированная группа шумов также подразделяется на световые и темновые. Первые - неоднородность фоточувствительности (photo response non-uniformity - PRNU), они описывают различие размеров эффективной части пикселя. Вторые – неоднородность темнового сигнала (dark signal non-uniformity - DSNU), они описывают чувствительность сенсора к темновым токам и случайному рождению электронов в ячейках пикселя. Детерминированные шумы также называют пространственными, так как различия сигналов при их наличии возникают при рассмотрении пикселей всего фотосенсора. Для математического описания шумов цифровых камер применяют среднеквадратические отклонения (СКО):

$$\sigma_{tot}(S) = \sqrt{\sigma_{LTN}^2 + \sigma_{DTN}^2 + \sigma_{LSN}^2 + \sigma_{DSN}^2} = \sqrt{S \cdot K + \sigma_{DTN}^2 + (S \cdot PRNU)^2 + DSNU^2}, \quad (1)$$

где, σ_{tot} – суммарное СКО шумов цифровой камеры, σ_{LTN} – СКО световых временных шумов, σ_{DTN} – СКО темновых временных шумов, σ_{LSN} – СКО световых пространственных шумов (цифр. ед.), σ_{DSN} – СКО темновых пространственных шумов (цифр. ед.), S – сигнал в цифровых единицах (цифр. ед.), K – коэффициент пересчёта электронов в цифровые единицы [эл./цифр.ед.].

В данном случае все СКО шумов выражены в цифровых единицах – условной величине, получаемой после обработки сигнала в электронах, накопленного в ячейке пикселя.

Световой временной шум описывается с помощью коэффициента пересчёта электронов в цифровые единицы K [эл./цифр.ед.]. Величина PRNU является безразмерной и определяется в долях или процентах.

Метод определения шумов по стандарту EMVA – референтный

В данной работе стандартный метод EMVA будет использоваться в качестве референтного. Для его реализации необходимо зарегистрировать большое количество снимков равномерной сцены на различных уровнях сигнала. Критерий равномерности сцены по EMVA представляется в виде:

$$\Delta S = \frac{S_{max} - S_{min}}{S_{mean}} \cdot 100\ \% < 3\%, \quad (2)$$

где - ΔS – максимальная разность сигналов однородного снимка, нормированная на среднее арифметическое, выраженное в процентах, S_{max} – максимальный сигнал кадра, S_{min} – минимальный сигнал кадра, S_{mean} – среднее арифметическое всех сигналов кадра.

Рекомендуется использовать не менее 50 различных уровней яркости. Среди них обязательно должны присутствовать минимальный (снимки с минимальной экспозицией при закрытой крышке камеры) и срединный уровень яркости. Для каждого уровня яркости требуется как минимум 2 снимка. Снимки на различных яркостях необходимы для построения зависимости временного шума от уровня сигнала, серия снимков на минимальном уровне сигнала требуется для расчёта темновых составляющих шума. Серия снимков на срединном уровне яркости требуется для корректного определения коэффициента PRNU. Так, используя набор данных из сотни (или более) изображений можно определить все составляющие шумов цифровых камер.

Чтобы осуществить обработку зарегистрированных серий снимков и рассчитать все шумовые составляющие в соответствии с EMVA следует использовать выражения (случай 2 снимков):

$$\sigma_{DTN}^2 = E\{Var\{D\}_t\}_{xy}, \quad (3)$$

$$DSNU^2 = Var\{E\{D\}_t\}_{xy} - \sigma_{DTN}^2, \quad (4)$$

$$\sigma_{LTN}^2 = E\{Var\{S\}_t\}_{xy} - \sigma_{DTN}^2, \quad (5)$$

$$PRNU^2 = (Var\{E\{S\}_t\}_{xy} - \sigma_{LTN}^2 - DSNU^2 - \sigma_{DTN}^2)/(E\{S\}_{xy} - E\{D\}_{xy}), \quad (6)$$

где E{●}j и Var{●}j – операторы математического ожидания и дисперсии соответственно, j – обозначает область пространства координат, по которым проводится операция, $D(D_{xyt})$ – трёхмерный массив темновых снимков в цифр. ед., x и y обозначают пространство координат в плоскости изображения, t обозначает координату времени (номера снимка) , $S(S_{xyt})$ – трёхмерный массив световых снимков в цифр. ед.

Метод оценки шума по одному снимку

В работе [2] применяют пуассон-гауссову модель шума, где одно пуассоново распределение соответствует фотонному дробовому шуму, а другое гауссово - остальным составляющим. Модель образования сигнала подразумевает сумму 2 составляющих: не зашумлённого приходящего сигнала и слагаемого, соответсвуюещго шуму. Второе слагаемое является произведением стандартного отклонения всех компонент шума на независимый случайный фактор с нулевым средним и СКО равным единице. Последнее в свою очередь принимается как сумма пуассоновской и гауссовой составляющих шума. Пуассонова составляющая аппроксимируется нормальным распределением при больших значениях математического ожидания, что и используется для упрощения выражений.

Целью является оценка стандартного отклонения всех составляющих шума. Конечным результатом применения метода является график зависимости стандартного отклонения от среднего значения сигналов по группам пикселей. Принципиальный алгоритм работы метода:

1) Получение двух матриц с вейвлет коэффициентами.

2) Разбиение матриц вейвлет коэффициентов по группам относительно сигналов в пикселях.
3) Расчёт средних значения и значений стандартного отклонения для каждого набора коэффициентов (сегментация), сформированного на шаге (2).

Предлагаемый метод получения шумовых параметров в соответствии с EMVA 1288

Для того, чтобы получить корректные значения шумовых параметров из данных в результате применения метода оценки шумов по одному изображению предлагается использовать взвешенное усреднение параметров шума, полученных по множеству оценок.

1) Считывание массива сигналов и массива отклонений;
2) Удаление данных, содержащих сигналы менее 10% от максимального сигнала для данного снимка и более 90%
3) Разбиение полученной зависимости шума от уровня сигнала на участки:
1. Нахождение разности минимального и максимального сигналов;
2. Формирование массива из чисел, полученных делением разности минимального и максимального сигнала на числа от 10 до 100000 с интервалом в порядок (т.е. всего 5)
3. Выбор числа из массива, находящегося в интервале от 10 до 100
4. Расчёт шага для разбиения зависимости шума от уровня сигнала делением разности минимального и максимального сигналов на число, полученное на шаге 3.
4) Использование точек на зависимости, разделённых интервалами, рассчитанными на шаге 4 как центров отсчёта новых интервалов. Малый шаг отсчёта новых интервалов составляет половину от полученного ранее. Откладывая малые интервалы (по одному за новый шаг) в сторону уменьшения и увеличения сигнала зависимость разделяется на участки различного размера. Действия на одном большом изначальном шаге заканчиваются при достижении концами интервала минимума и максимума.
5) Аппроксимация зависимости общего шума от уровня сигнала на каждом интервале в соответствии с выражением $\sqrt{a^2x^2 + bx + c^2}$ (где x сигнал)с параметрами a (PRNU), b (K) и c ($\sqrt{\sigma_{dt}^2 + DSNU^2}$).
   1. Аппроксимация зависимости временного шума от уровня сигнала на каждом интервале в соответствии с выражением $\sqrt{a_1x + b_1^2}$ с параметрами a_1 (K) и b_1 ($\sqrt{\sigma_{dt}^2 + DSNU^2}$).
   2. Аппроксимация зависимости пространственного шума от уровня сигнала в соответствии с выражением $\sqrt{a_2^2x^2 + b_2^2}$ с параметрами a_2 (PRNU) и b_2 ($\sqrt{\sigma_{dt}^2 + DSNU^2}$).
   3. Использование a_1 как начального значения параметра a, a_2 как начального значения параметра b и наименьшего из b1 и b2 как начального значения параметра c при аппроксимации зависимости общего шума от уровня сигнала
6) Формирование массивов оценок PRNU, K, и темновых шумов, полученных из аппроксимаций по всем описанным интервалам. Одновременное формирование соответствующих оценкам массивов длин рассматриваемых интервалов и СКО (Важно отметить, что СКО получается при расчёте отклонения зависимости, рассчитанной на основе анализа интервала от данных во всём диапазоне сигналов)
7) Расчёт весовых коэффициентов для PRNU, K, и темновых шумов. Каждый из весовых коэффициентов имеет два сомножителя, первый сомножитель является общим для всех 3 весов и рассчитывается как отношение нормированной величины интервала рассматриваемой оценки к нормированному СКО для данной оценки. Уникальные сомножители для каждого из параметров формировались на основе экспериментальных наблюдений внутренней статистики массивов оценок PRNU, K, и темновых шумов:
   PRNU) В EMVA 1288 упоминается, что PRNU имеет величину порядка 10^{-3}, что подтверждается экспериментальными наблюдениями. Вес для PRNU рассчитывается как модуль нормированной разности десятичных логарифмов PRNU (для данного интервала)

и 10^{-3}. Чтобы обеспечить необходимое распределение весов, полученные величины вычитаются из 1.

K) В силу того, что оценка данной величины в большинстве случаев близка к получаемым более точными методами вес для K рассчитывается как модуль нормированной разности медианного и рассматриваемого значения. Чтобы обеспечить необходимое распределение весов, полученные величины вычитаются из 1.

Темновые шумы) В силу экспериментального наблюдения малости оценок близких к тем, что удавалось получить при использовании стандартного метода EMVA 1288 вес для темнового шума зависел от числа получаемых близких оценок – чем меньше таковых, тем больше вес.

8) Получение финальных оценок световых и темновых шумов при использовании стандартной формулы взвешенного среднего: $\frac{\sum W \times Y}{\sum W}$, где W соответствующий вес оценки Y.

Экспериментальная апробация предложенного метода

Чтобы сравнить рассматриваемые методы был проведен эксперимент по измерению шумов цифровых камер. В качестве источника излучения использовался лазер. Чтобы сделать лазерный пучок однородным, использовались линза (микрообъектив) и диафрагма. Чтобы снизить пространственную когерентность излучения применялся деспеклер. Съёмка проводилась при управлении камерой с компьютера. При помощи описанной установки проводились измерения шумовых характеристик монохромной ПЗС камеры для научных задач Retiga R6, которая имеет АЦП 14 бит, 5913600 пикселей, минимальное время экспозиции 0,0006 мс, размер пикселя 4.6 мкм,

График зависимости временного шума от уровня сигнала для Retiga R6, полученный по методу EMVA 1288 представлен на рисунке 1.

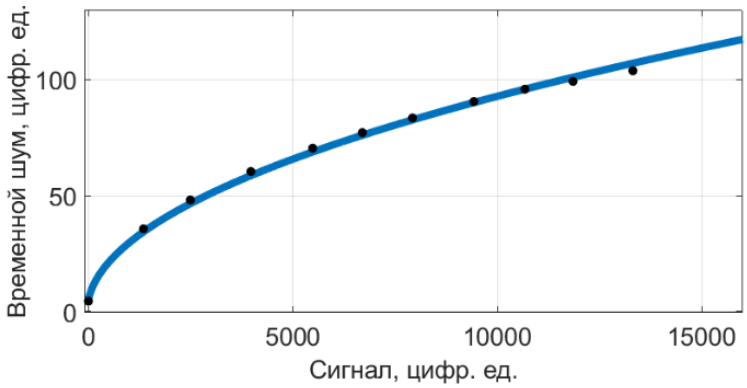
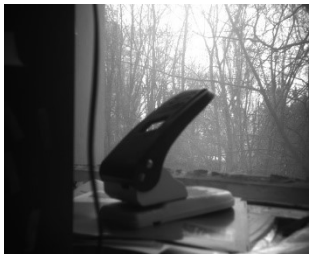


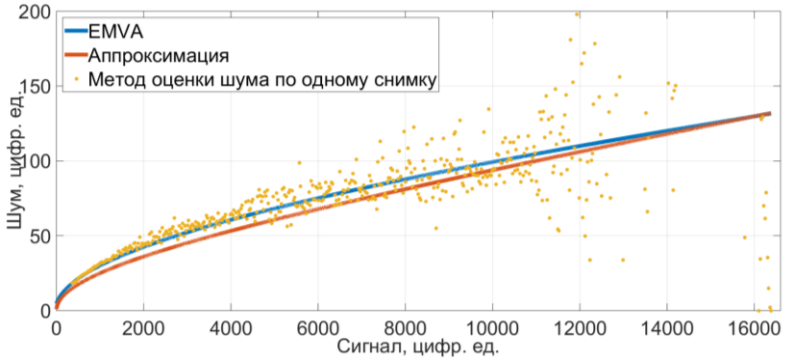
Рисунок 1 - Зависимость временного шума от уровня сигнала полученная по EMVA 1288

Получены следующие значения шумовых характеристик: K = 0.84±0.03, цифр.ед. / эл., PRNU = 0.0033±0.0001, отн.ед., DSNU = 0.50±0,01, цифр.ед., σ_{dt} = 4.46±0.01, цифр.ед. Таким образом, темновой шум составляет $\sqrt{DSNU^2 + \sigma_{dt}^2} = 4.49 \pm 0.01$.

Результаты применения разработанного алгоритма представлены в рисунке 2 и в таблице 1. Для сравнения в таблице 3 представлены результаты полученные при прямой аппроксимации.



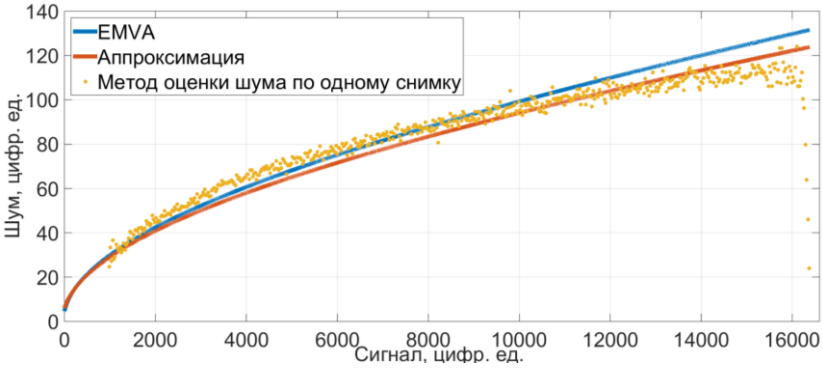
а



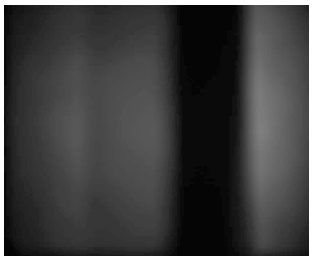
б



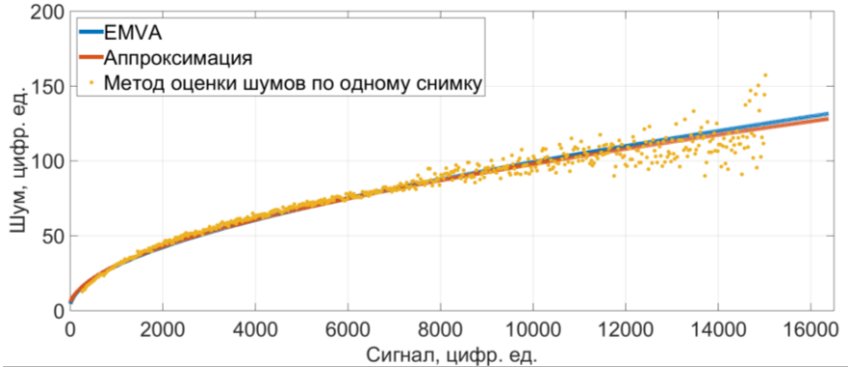
в



г



д



е

Рисунок 2 – Изображения (а, в, д) и полученные зависимости (б, г, е)

Таблица 1. Результаты оценки шумовых параметров предложенным методом (выделено жирным шрифтом) и при прямой аппроксимации

Характеристики	Тип сцены		
	Бытовая	Неоднородная	Полосовая
K, эл, цифр. ед	**0.593** / 0.974	**0.799** / 0.881	**0.890** / 0.976
PRNU, отн. ед.	**0.0054** / 0.0039	**0.0029** / 0.0027	**0.0026** / $1.5 \cdot 10^{-7}$
Темновой шум, цифр. ед.	**0.57** / 9.09	**5.89** / 23.29	**5.79** / $7.58 \cdot 10^{-11}$

Из полученных результатов видно, что метод оценки шумов по одному снимку позволяют получить достаточно достоверные данные для извлечения шумовых параметров цифровой камеры. Также, видно, что прямая аппроксимация данных, пример которой представлен в таблице 1 (нежирным) даёт разброс извлекаемых параметров в зависимости от сцены существенно больше, чем в случае применения предложенного алгоритма. Кроме того, видно,что чем более сцена близка к неоднородной, тем ближе результат оценки шумовых параметров к данным, полученным с помощью метода EMVA. Не смотря на то, что результаты оценок как по предложенному методу, так и при прямой аппроксимации сильно зависят от используемого изображения, однако предложенный метод позволяет получать оценки, совпадающие по порядку с оценками EMVA, а в случае использования достаточно неоднородной сцены погрешность по отношению к EMVA оказывается меньше 20%.

Литература

1. EMVA European Machine Vision Association, EMVA Standard 1288. Standard for characterization of image sensors and cameras 3.1 2016. Source: http://www.emva.org/cms/upload/ Standards/
2. Foi A. et al. Practical Poissonian-Gaussian noise modeling and fitting for single-image raw-data //IEEE transactions on image processing. 2008. Vol. 17. №. 10. P. 1737-1754.

Работа выполнена при финансовой поддержке Российского научного фонда (РНФ), грант № 24-19-00898.

ESTIMATION OF DIGITAL CAMERA NOISE CHARACTERISTICS FROM A SINGLE IMAGE

Kozlov A.V.[1,2]

[1] National research nuclear university MEPhI
[2] Russian Metrological Institute of Technical Physics and Radio Engineering

An algorithm for estimation of noise parameters of digital camera photosensors from the data obtained by applying the single image noise estimation algorithm is presented. The proposed method is tested on images of a prepared scene and on an image of an ordinary common household scene. It is obtained that the proposed method gives more accurate estimations in comparison with direct approximation of the data obtained as a result of application of the algorithm of single image noise estimation by an expression for the total standard deviation of all noise components.

ПРОГРАММНО-АППАРАТНЫЙ КОМПЛЕКС ДЛЯ ИЗМЕРЕНИЯ ДИЭЛЕКТРИЧЕСКОЙ ПРОНИЦАЕМОСТИ МАТЕРИАЛОВ С ИСПОЛЬЗОВАНИЕМ МЕТОДА НИКОЛСОНА–РОССА–ВЕЙРА

зам. нач. отдела 11 Бондаренко А.С. [1], с.н.с., к.т.н. Боровков А.С. [1,2], инж. Жуков Д.А. [1], вед. инж. Раков А.В. [1], нач. лаб.113 Семёнов В.А. [1]

[1] ФГУП «ВНИИФТРИ», Менделеево, Московская обл.
[2] Национальный исследовательский университет «МИЭТ», Москва

В работе рассмотрен вопрос измерения комплексной диэлектрической и магнитной проницаемостей материалов объёмных поглотителей нагрузок волноводных. При решении измерительной задачи использовались результаты исследований, полученные после применения модифицированного метода Николсона-Росса-Вейра. Разработана программа для расчёта комплексной диэлектрической и магнитной проницаемостей по результатам измерений комплексных коэффициентов отражения и передачи волноводного отрезка линии передачи, в который помещён исследуемый образец материала с известными геометрическими размерами. Проведена оценка составляющих погрешностей комплексной диэлектрической и магнитной проницаемостей, обусловленных погрешностями измерений комплексных коэффициентов отражения и передачи в волноводных трактах. Для реализации данной процедуры использовался метод Монте-Карло. Расчёты проводились методом цифровой обработки S-параметров матриц рассеяния четырёхполюсников. Установлено, что для данного метода при применении рабочих эталонов единиц комплексных коэффициентов отражения и передачи исследуемые составляющие погрешности измерений комплексных диэлектрической и магнитной проницаемостей составляют 2 – 3 %. Проведено сравнение результатов расчётов с помощью разработанной программы с результатами моделирования в программе электродинамического моделирования CST. Показана сходимость результатов моделирования и экспериментов, что подтверждает достоверность расчётов, проводимых с помощью разработанной программы. Показана целесообразность применения программы при разработке и исследованиях объёмных поглотителей электромагнитной энергии, применяемых в СВЧ-устройствах и узлах.

Введение

Измерение комплексной диэлектрической и магнитной проницаемостей материалов является актуальной задачей при создании современных устройств СВЧ электроники. К таким устройствам в том числе относятся согласованные волноводные нагрузки с объёмным поглотителем. Величины комплексных диэлектрической и магнитной проницаемостей материала поглотителя (ε_m и μ_m) совместно с геометрическими размерами линии передачи определяют волновое сопротивление нагрузки и, следовательно, величину комплексного коэффициента отражения (ККО). Проверка на соответствие номинальным значениям ККО при производстве объёмных поглотителей нагрузок проводится с помощью установок для измерений ε_m и μ_m. Измеренные значения ε_m и μ_m являются ключевыми показателями при принятии решения о соответствии поглотителя заданным требованиям. С уменьшением сечения волноводной и коаксиальной линии требуется большая точность измерений данных характеристик. Существует несколько методов определения ε_m и μ_m. Каждый из них имеет свои преимущества и ограничения и обеспечивает разную точность измерений. Проверка соответствия метода измерений исследуемому физическому объекту является важной и актуальной задачей.

В данной статье представлены результаты применения разработанной программы, в которой был реализован метод Николсона-Росса-Вейра [1 – 3] для расчета ε_m и μ_m материала объёмного поглотителя согласованной нагрузки. Программа разработана в среде программирования Mathlab. Суть метода заключается в определении ε_m и μ_m на основе результатов измерений S-параметров образца материала, помещенного в волноводный или коаксиальный отрезок линии передачи. Проведены экспериментальные исследования данного метода. Проверка результатов измерений с использованием разработанной программы проводилась путём сравнения с результатами моделирования в программе CST.

Комплексная диэлектрическая и магнитная проницаемости объёмных поглотителей

Комплексные диэлектрическая и магнитная проницаемости диэлектрика играют ключевую роль в описании электромагнитных свойств материалов. Их выражают с помощью формул

$$\varepsilon_m = \varepsilon' + j\varepsilon'', \qquad \mu_m = \mu' + j\mu''. \tag{1}$$

Действительная часть диэлектрической проницаемости указывает на способность материала к поляризации под воздействием внешнего электрического поля. Мнимая составляющая связана с диэлектрическими потерями, которые включают в себя потери энергии из-за перемещения зарядов. Действительная часть комплексной магнитной проницаемости является показателем способности материала намагничиваться в ответ на приложенное магнитное поле, мнимая составляющая – показатель магнитных потерь, которые возникают из-за гистерезиса, вихревых токов и других факторов.

Исследуемый объект в данной работе – объёмный поглотитель (рисунок 1) согласованной нагрузки, изготовленный из ферроэпоксида, который имеет вид треугольной пирамиды. Такая конфигурация рассчитана на поглощение магнитной составляющей электромагнитной энергии. Главным параметром ферроэпоксидного объёмного поглотителя является мнимая часть магнитной проницаемости материала. Также иногда на практике используются объёмные поглотители из проводящих материалов или диэлектриков, имеющие другие геометрические формы. Такие поглотители рассчитаны на погашение электрической составляющей электромагнитной энергии, в связи с чем, он имеет форму правильной пирамиды или клина.

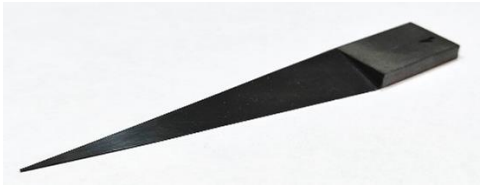


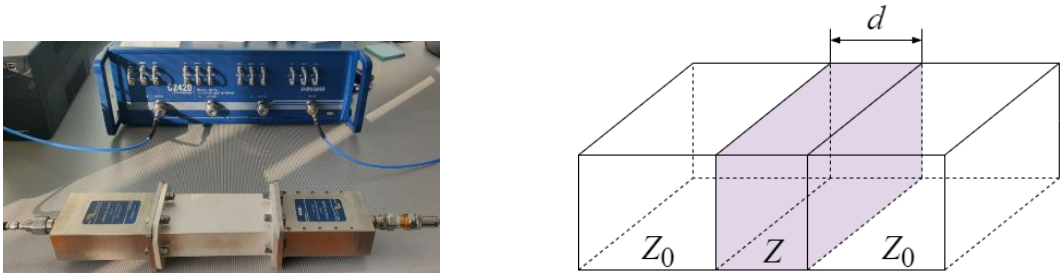
Рисунок 1 – Объёмный ферроэпоксидный поглотитель согласованной нагрузки

Известны методы измерений ε_m и μ_m объемных поглотителей с использованием коаксиального пробника или линии передачи, метод измерений в свободном пространстве, методы, в которых применяются объемные резонаторы или параллельные пластины. Резонансный метод наиболее часто применяется при измерениях ε_m и μ_m. Этот метод обладает рядом недостатков. Например, требуемые величины могут быть получены только на особых резонансных частотных точках, что не всегда является достаточным. Иногда требуется знать данные значения во всём диапазоне частот. Кроме того, резонанс возможен только в образцах с малыми потерями. Большой широкополосностью обладают методы, основанные на измерениях в линии передачи. Одним из методов определения ε_m и μ_m является модифицированный метод Николсона-Росса-Вейра [4, 5], который был взят за основу исследования настоящей работы. В данном методе образец материала представлен в виде четырехполюсника, расчёт комплексных диэлектрической и магнитной проницаемостей проводится с использованием измеренных S-параметров материала, помещённого в отрезок коаксиальной или волноводной линии передачи.

Описание эксперимента

В рамках данной работы было проведено экспериментальное исследование образца, изготовленного из ферроэпоксида, применяемого в качестве поглощающего элемента согласованных волноводных нагрузок. Измерения проводились с использованием отрезка волноводной линии передачи с размерами сечения 58 х 25 мм. Образцу была предана форма прямоугольной пластины известной длины, а его ширина и высота равны размерам широкой и узкой стенок волновода. Измерительная установка для проведения эксперимента по определению ε_m и μ_m диэлектрического образца включает в себя векторный анализатор цепей (ВАЦ) производства «Планар» и кабельную сборку, состоящую из коаксиальных кабелей, волноводно-коаксиальных переходов и волноводного отрезок линии передачи. Калибровка ВАЦ проводилась по алгоритму метода TRL (от англ. Thru – соединение двух волноводных портов, Reflect – мера «короткое замыкание» и Line – отрезок прямоугольного волновода). Вначале были измерены ККО и комплексные коэффициенты передачи (ККП) полой волноводной секции. Далее испытуемый образец разместили в волноводной секции на известном расстоянии от первого порта и повторили измерения (рисунок 2 а).

Согласно модели измерений образец материала в виде пластины известной толщины (d) был размещен внутри прямоугольного волновода с характеристическим импедансом Z_0 (рисунок 2 б). Геометрические размеры образца совпадают с размерами широкой и узкой стенок волноводного отрезка линии. Электромагнитные волны распространяются вдоль прямоугольного волноводного отрезка линии передачи.



а) б)

Рисунок 2 – а) Установка для измерений параметров материала,
б) Модель отрезка прямоугольного волновода с размещённым в нём образцом

S-параметры волноводного отрезка линии прямоугольного сечения с размещённым в нём образцом определяются в соответствии с формулами

$$S_{11} = S_{11\text{m}}\left(1-S_{21\text{m}}^2\right)/\left(1-S_{11\text{m}}^2S_{21\text{m}}^2\right), \qquad S_{21} = S_{21\text{m}}\left(1-S_{11\text{m}}^2\right)/\left(1-S_{11\text{m}}^2S_{21\text{m}}^2\right),$$

$$S_{11\text{m}} = \left(Z-Z_0\right)/\left(Z+Z_0\right),$$

$$Z = \text{j}\omega\mu_0\mu_\text{m}/\gamma, \qquad Z_0 = \text{j}\omega\mu_0/\gamma_0 \tag{2}$$

где $S_{11\text{m}}$ – коэффициент отражения образца в плоскости раздела сред с волновыми сопротивлениями Z и Z_0, $S_{21\text{m}}$ – коэффициент передачи образца конечной длины, который определяется по формуле

$$S_{21\text{m}} = \text{e}^{-\gamma d},$$

γ – постоянная распространения для линии передачи с размещённым в ней образцом

$$\gamma = \text{j}\left(2\pi/\lambda_0\right)\sqrt{\varepsilon_\text{m}\mu_\text{m}-\left(\lambda_0/\lambda_\text{c}\right)^2}, \tag{3}$$

γ_0 для линии передачи без образца определяется выражением

$$\gamma_0 = \text{j}\left(2\pi/\lambda_0\right)\sqrt{1-\left(\lambda_0/\lambda_\text{c}\right)^2}, \tag{4}$$

λ_0 – длина волны в свободном пространстве, λ_c – критическая длина волны.
Комплексные коэффициенты $S_{11\text{m}}$ и $S_{21\text{m}}$ могут быть рассчитаны по формулам

$$\begin{cases} S_{11\text{m}} = K+\sqrt{K^2-1}, & \text{при } \left|K+\sqrt{K^2-1}\right| \le 1, \\ S_{11\text{m}} = K-\sqrt{K^2-1} & \text{при } \left|K+\sqrt{K^2-1}\right| > 1, \end{cases}$$

$$K = \left(S_{11}^2-S_{21}^2+1\right)/\left(2S_{11}\right),$$

$$S_{21\text{m}} = \frac{S_{11}+S_{21}-S_{11\text{m}}}{1-S_{11\text{m}}\left(S_{11}+S_{21}\right)}.$$

Учитывая, что $Z/Z_0=(1+S_{11\text{m}})/(1-S_{11\text{m}})$ и подставив (3) и (4) в (2), формула комплексной магнитной проницаемости приобретёт вид

$$\mu_\text{m} = \frac{\lambda_\text{w}}{\lambda_\text{m}}\left(\frac{1-S_{11\text{m}}}{1+S_{11\text{m}}}\right), \tag{5}$$

$$\lambda_\text{w} = \lambda_0/\sqrt{1-\left(\lambda_0/\lambda_c\right)^2}, \qquad \lambda_\text{m} = \lambda_0/\sqrt{\varepsilon_\text{m}\mu_\text{m}-\left(\lambda_0/\lambda_\text{c}\right)^2}, \tag{6}$$

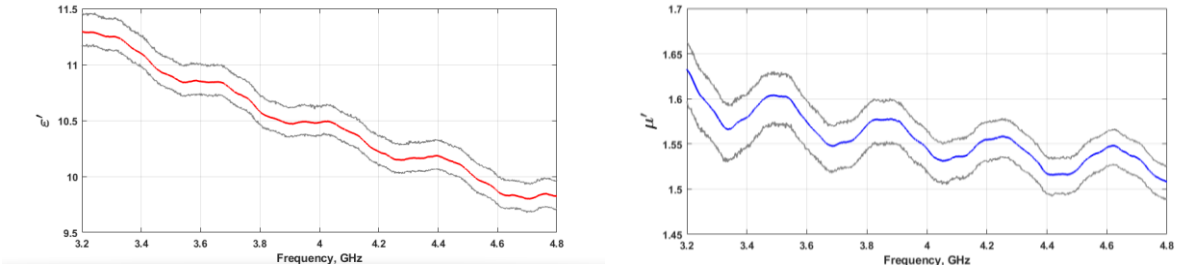
где λ_w – длина волны в волноводе без образца, а λ_m – длина волны в волноводе с размещённым в нём образцом. После подстановки (6) в (5) и преобразования получим

$$\varepsilon_\text{m} = \lambda_0^2\left(\lambda_\text{m}^{-2}+\lambda_\text{c}^{-1}\right)/\mu_\text{m}.$$

Формулы использовались в дальнейшем при разработке программы для определения ε_m и μ_m.

Исследование составляющей погрешности измерений ε_m и μ_m, обусловленных погрешностями измерений ККО и ККП в волноводных трактах проводилось с использованием метода Монте-Карло. Суть метода состоит в том, что для математической модели измерений отклонение входных величин в границах их погрешностей задаются с использованием генератора случайных чисел, проводится многократный расчет с использованием модели, погрешности выходных величин определяются по их разбросам, полученным в результате многократных расчётов модели. Входные параметры для программы: значения ККО и ККП волноводной секции с материалом и без него, геометрические размеры тракта (сечение и длина) и испытуемого образца, расстояние материала от первого порта, а также пределы допускаемой погрешности ККО и ККП, выходные параметры – ε_m и μ_m с их пределами погрешности.

При расчёте с помощью программы определены значения ε_m и μ_m и их погрешности. На графиках представлены частотные зависимости ε_m и μ_m с предельными отклонениями для заданных значений показателей точности рабочих эталонов единиц ККО и ККП в волноводных трактах.



а) б)

Рисунок 4 – Значения действительных частей комплексных а) диэлектрической проницаемости, б) магнитной проницаемости

Моделирование поглотителя

Достоверность расчётов, проводимых при моделировании, зависит от степени соответствия планируемому эксперименту и применения известных функциональных возможностей программы CST Microwave Studio. Поэтому вначале была разработана модель и рассчитаны S-параметры для волноводного отрезка линии передачи с размерами сечения 7,2 х 3,4 мм. Далее проводилось моделирование образца в виде прямоугольного параллелепипеда с заданными ему ε, μ и известными геометрическими размерами. Заданы значения: $\varepsilon = 5$; тангенс угла диэлектрических потерь $\tan\delta(\varepsilon) = 0{,}002$; $\mu = 1$; $\tan\delta(\mu) = 0{,}04$, начальная частота $f_{min} = 25{,}95$ ГГц, конечная частота $f_{max} = 37{,}5$ ГГц. Модель образца была размещена в волноводной секции (рисунок 3). Расчёс S-параметров для конечной модели проводился в режиме Frequency Solver и автоматических настройках в разделе Global Properties.

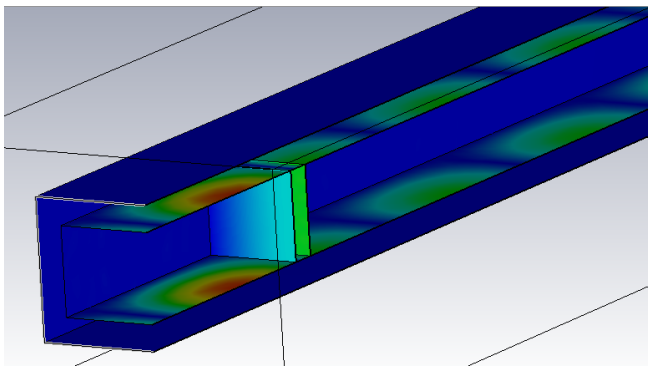


Рисунок 3 – Модель волноводной линии с исследуемым образцом

Первым этапом разработки модели и программы расчета ε_m и μ_m является определение граничных условий и требований к исследуемому образцу и измерительной установке [6]. Они необходимы для получения достоверных значений ККО и ККП, а также для проведения корректного сравнения результатов моделирования с результатами расчёта с помощью разработанной программы. Диэлектрик должен иметь такое же сечение или диаметры, как и измерительный тракт, а его длина не должна превышать половину длины волны, распространяемой в среде для того, чтобы при помещении образца в волноводную секцию была возможность отодвинуть материал от порта на известное расстояние и не попасть в область пучностей электромагнитной волны. В противном случае конечный результат будет обладать высокой погрешностью. Измерительная секция также имеет требования, т.к. она обладает потерями. Требуется учитывать её влияние на конечные исследуемые S-параметры материала для того, чтобы получить достоверные результаты измерений образца. Для этого волноводная секция должна быть существенной длины, порядка нескольких длин полуволн. От этого напрямую зависит точность конечных результатов.

Проверка достоверности результатов расчётов с помощью программы осуществлялась по результатам моделирования в программе CST Microwave Studio. Touchstone (s2p) файлы, полученные при моделировании, были использованы при расчёте ε_m и μ_m с помощью разработанной программы. На графиках рисунка 4 представлены частотные зависимости действительных частей ε_m и μ_m, заданных при моделировании (ε'_1 и μ'_1) и полученных с помощью разработанной программы (ε'_2 и μ'_2).

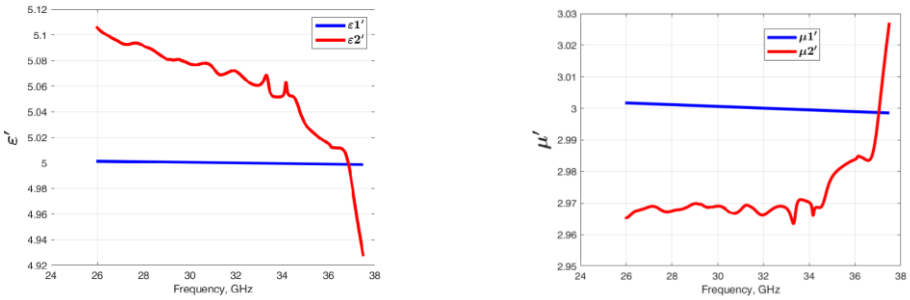


Рисунок 4 – Действительные части ε_m и μ_m, (ε'_1 и μ'_1 – заданы при моделировании, ε'_2 и μ'_2 – получены с помощью программы)

Как видно из графиков разности между ε'_1 и ε'_2 составляют не более 0,1, а между μ'_1 и μ'_2 – 0,04, что не превышает значений исследуемых погрешностей.

Результаты работы

По результатам проведённой работы разработана программа по расчёту ε_m и μ_m с оценкой составляющей погрешности, обусловленнной показателями точности измерительной установки. Установлено, что при реализации метода Николсона-Росса-Вейра с использованием в измерительных установках рабочих эталонов единиц ККО и ККП исследуемые составляющие погрешности измерений ε_m и μ_m не превышают 2 – 3 %. В результате моделирования эксперимента с образцом с заданными ε_m и μ_m получены Touchstone (s2p) файлы. Рассчитанные по этим данным значения ε_m и μ_m соответствуют в заданных пределах значениям, установленным при моделировании в программе CST Microwave Studio, что подтверждает достоверность полученных расчётов с помощью разработанной программы. Показана целесообразность применения программы при разработке и исследованиях объёмных поглотителей электромагнитной энергии, применяемых в СВЧ-устройствах и узлах. Полученные результаты могут быть использованы для уточнения рецептуры при изготовлении объёмных поглотителей согласованных волноводных нагрузок.

Литература

1. Nicolson A. M., Ross G. F. Measurement of the intrinsic properties of materials by time-domain techniques //IEEE Transactions on instrumentation and measurement. – 1970. – Т. 19. – №. 4. – С. 377-382.
2. Rothwell E. J. et al. Analysis of the Nicolson-Ross-Weir method for characterizing the electromagnetic properties of engineered materials //Progress In Electromagnetics Research. – 2016. – Т. 157. – С. 31-47.
3. Antyasov I. S., Sokolov A. N. Investigation limits of applicability of the method Nicolson-Ross-Weir to calculate the permittivity and permeability radar absorbing material //2016 2nd International Conference on Industrial Engineering, Applications and Manufacturing (ICIEAM). – IEEE, 2016. – С. 1–5.
4. de Paula A. L., Rezende M. C., Barroso J. J. Modified Nicolson-Ross-Weir (NRW) method to retrieve the constitutive parameters of low-loss materials //2011 SBMO/IEEE MTT-S international microwave and optoelectronics conference (IMOC 2011). – IEEE, 2011. – С. 488-492.
5. Huashen W. et al. Electromagnetic parameters test system based on a refined NRW transmission/reflection algorithm //2007 International Symposium on Microwave, Antenna, Propagation and EMC Technologies for Wireless Communications. – IEEE, 2007. – С. 1276-1280.
6. Vicente A. N., Dip G. M., Junqueira C. The step by step development of NRW method //2011 SBMO/IEEE MTT-S International Microwave and Optoelectronics Conference (IMOC 2011). – IEEE, 2011. – С. 738-742.

HARDWARE AND SOFTWARE COMPLEX FOR MEASURING THE DIELECTRIC PERMITTIVITY OF MATERIALS USING THE NICHOLSON-ROSS-WEIR METHOD

Bondarenko A.S.[1], Borovkov A. S.[1,2], Zhukov D. A.[1], Rakov A.V.[1], Semyonov V. A.[1]

[1] Russian metrological institute of technical physics and radio engineering
Russia, Moscow region, Mendeleevo
[2] National Research University of Electronic Technology
Russia, Moscow, Zelenograd, Shokin square, 1

The paper considers the issue of measuring the complex permittivity and magnetic permeability of materials of volume absorbers of waveguide loads. When solving the measurement problem, the results of studies obtained after applying the modified Nicholson-Ross-Weir method were used. A program has been developed for calculating the complex permittivity and magnetic permeability based on the results of measuring the complex reflection and transmission coefficients of a waveguide section of a transmission line, in which the studied sample of the material with known geometric dimensions is placed. An assessment has been made of the error components of the complex permittivity and magnetic permeability caused by the measurement errors of the complex reflection and transmission coefficients in waveguide paths. The Monte Carlo method was used to implement this procedure. The calculations were carried out using the digital processing of the S-parameters of the scattering matrices of four-terminal networks. It has been found that for this method, when using working standards of units of complex reflection and transmission coefficients, the studied components of the measurement error of the complex permittivity and magnetic permeability are 2 - 3%. The results of calculations using the developed program are compared with the results of modeling in the CST electrodynamic modeling program. The convergence of the results of modeling and experiments is shown, which confirms the reliability of the calculations carried out using the developed program. The expediency of using the program in the development and research of volumetric absorbers of electromagnetic energy used in microwave devices and units is shown.

27-я Международная конференция «Цифровая обработка сигналов и её применение – DSPA-2025», Москва, Россия, доклады.

Серия: Цифровая обработка сигналов и её применение

Выпуск: XXVII

Почтовый адрес:

107031, Москва, ул. Рождественка, д. 6/9/20

Исполнительная Дирекция РНТОРЭС им. А.С. Попова

E-mail организационного комитета конференции:
dspa@dspa-conf.org

Группа в Telegram:
https://t.me/+IFIWdpcBCO4yMTdi

Веб-сайт конференции: www.dspa-conf.org

Тел. РНТОРЭС им. Попова: +7 (903) 201-53-33